ORIGINAL ARTICLE

Veterinary Research Forum. 2021; 12 (2) 157 – 165
doi: 10.30466/vrf.2019.98625.2351
Journal Homepage: vrf.iranjournals.ir

Veterinary Research Forum

In silico prediction of B and T cell epitopes based on NDV fusion protein for vaccine development against Newcastle disease virus

Seyedeh Somayeh Hosseini[1], Khosrow Aghaiypour Kolyani[2]*, Robab Rafiei Tabatabaei[3], Hossein Goudarzi[4], Abbas Akhavan Sepahi[5], Maryam Salemi[2]

[1] Department of Microbiology, Islamic Azad University, Tehran North Branch, Tehran, Iran; [2] Department of Genomics and Genetic Engineering, Razi Vaccine and Serum Research Institute (RVSRI), Agricultural Research, Education and Extension Organization (AREEO), Karaj, Iran; [3] Department of Microbiology, Faculty of Basic Sciences, Islamic Azad University, Tehran North Branch, Tehran, Iran; [4] Central Laboratory Department, Razi Vaccine and Serum Research Institute Agricultural Research, AREEO, Karaj, Iran; [5] Department of Microbiology, Faculty of Science, Tehran North Branch, Islamic Azad University, Tehran, Iran.

Article Info

Article history:

Received: 29 November 2018
Accepted: 07 May 2019
Available online: 15 June 2021

Keywords:

Epitope
Docking
Newcastle disease virus
Peptide-based vaccine

Abstract

Newcastle disease (ND) is known as the most common diseases of economic importance worldwide. Vaccination against virulent strains of Newcastle disease virus (NDV) has failed during some outbreaks. Here, we aimed to assess the epitopes of NDV fusion protein as targets for a peptide-based vaccine. To explore the most antigenic epitopes on the F protein, we retrieved virulent strains of genotype VII from National Center for Biotechnology Information (NCBI). Linear and conformational B-cell epitopes were identified. Moreover, T-cell epitopes with high and moderate binding affinities to human major histocompatibility complex (MHC) class I and class II alleles were predicted using bioinformatics tools. Subsequently, the overlapped epitopes of B-cell and MHC class I and MHC class II were determined. To validate our predictions, the best epitopes were docked, to chicken MHC class I (B-F) alleles using the HADDOCK flexible docking server. Seven 'high ranked epitopes' were identified. Among them, 'LYCTRIVTF' and 'MRATYLETL' showed the highest scores. The other five epitopes including LSGEFDATY, LTTPPYMALK, LYLTELTTV, DCIKITQQV and SIAATNEAV obtained very encouraging results as well. SIAATNEAV had been recognized as a neutralizing epitope of F protein using monoclonal antibodies before. Taken together, our results demonstrated that the identified epitopes needed to be tested by *in vitro* and *in vivo* experiments.

Introduction

Being a member of family *Paramyxoviridae*, Newcastle disease virus (NDV) contains single stranded RNA (approximately 15kb). NDV is commonly known as avian paramyxovirus serotype 1 (APMV-1) and its virulent strains (vNDV) cause Newcastle disease (ND) in a wide range of avian species. The NDV along with avian influenza virus (AIV) are the two most serious avian pathogens which are intercontinentally distributed. ND gives rise to devastating economic losses in poultry industry globally.[1] NDV genome codes for six proteins including nucleo-protein (NP), phosphoprotein (P), matrix protein (M), fusion protein (F), surface glycoprotein hemagglutinin-neuraminidase (HN), and large RNA dependent RNA polymerase (L). Moreover, two additional proteins can be produced *via* RNA editing of the P protein.[2] According to sequence analysis of F gene, three classification systems have been introduced so far[3-5] In recent classification, NDV strains have been divided into class I and class II. While class I comprises only a single genotype, class II includes more than 18 genotypes of both low and high virulence.[2,6] All four panzootics of ND since 1920s have been caused by isolates of class II.[5] Among genotypes of class II, genotype VII has been responsible for the fourth panzootic, started in 1985 in Far East and still ongoing, and it has been isolated in Asia, Africa, Western Europe and even in South America. Viruses of sub-genotype VIId are of great importance as they are among the most prevalent NDV genotypes and are likely to spread to wild birds.[7,8]

***Correspondence:**

Khosrow Aghaiypour Kolyani. MSc, PhD
Department of Genomics and Genetic Engineering, Razi Vaccine and Serum Research Institute (RVSRI), Agricultural Research, Education and Extension Organization (AREEO), Karaj, Iran
E-mail: khosrow@rvsri.ac.ir

Vaccination programs against NDV are implemented in some countries, especially those with endemic NDV. Classical live or inactivated NDV vaccines are formulated with genotype I and/or II (low virulent NDV strains). In spite of using such vaccines, NDV outbreaks (notably isolates of genotype VII) are still reported globally even in vaccinated poultry flocks.[7,9] Currently, advances in computational approaches of vaccine design and availability of huge sequence information have attracted many researchers.[10] Using immunoinformatics approaches reduce the time and cost of vaccine development.[11] However, in case of NDV, not much research has been performed.[12,13] The NDV possesses two glycoproteins forming surface projections. Hemagglutinin-neuraminidase (HN) protein and has sialic acid binding sites responsible for virus attachment. Fusion (F) protein is involved in fusion of the virus with host cell. They are both capable of inducing neutralizing antibodies, however, the homologous F protein is shown to be of greater importance in conferring protection.[14] Many studies have shown the physico-chemical properties of F protein and these properties could be computed using an online tool at Expasy (http://web.expasy.org/protparam/).[14,15] The F protein is consisted of Leucine, Isoleucine, Alanine, Threonine, Glycine, Serine and Valine, and can be high-abundance amino acids represented about 61.00% of amino acid content of protein.

In the present work, we assessed the F protein of virulent NDV strains (genotype VII) *in silico* analysis to determine protective epitopes which paves the way for developing a peptide-based vaccine against NDV.

Materials and Methods

Fusion protein sequence retrieval and detection of conserved regions. A total number of 126 fusion protein sequences belonged to different sub-genotypes of VII in Asia, especially VIId circulating in Iran, were retrieved from NCBI database.[9,16] We sequenced an isolate of sub-genotype VIId from Iran, I was submitted to GenBank® (Accession number: KP347437) and selected as reference sequence in this study. To determine conserved regions, sequences were aligned through multiple sequence alignment using BioEdit software (version 7.1.9 Isis Pharmaceuticals, Carlsbad, USA).

Sequence and structure analysis of F protein. Physico-chemical properties of F protein was computed using an online tool at Expasy (http://web.expasy.org/protparam/). Leucine, isoleucine, alanine, threonine, glycine, serine and valine were determined to be high-abundance amino acids representing about 61.00% of amino acid content of protein. InterProScan (http://www.ebi.ac.uk/Tools/InterProScan/) is a signature scanning software which was used to determine cytoplasmic, non-cytoplasmic and transmembrane domain of the protein. Additionally, reference sequence was submitted to TMHMM server, which is used to predict the most probable topology and is relied on hidden Markov model (HMM). Secondary structure of protein was investigated using PSIPRED (http://bioinf.cs.ucl.ac.uk/psipred).

Prediction of B-cell epitopes. B-cell epitopes can be divided into two groups: Linear (continuous) and conformational (discontinuous) epitopes. Linear epitope is a short stretch of amino acids within a protein sequence while conformational epitope comprises distant residues of an antigen join from polypeptide folding. Despite containing short linear peptides, 90% of B-cell epitopes are conformational.[17]

Linear B-cell epitopes. BepiPred is known as a method which encompasses the hidden Markov model (http://tools.iedb.org/bcell). It is employed in pursuance of linear B-cell epitope identification with a threshold value of 1.00. BCPREDS is another tool for prediction which uses physico-chemical properties (http://ailab.ist.psu.edu/bcpred/predict.html).[18] To determine antigenicity of linear B-cell and T-cell epitopes, VaxiJen server was run with default parameters.

Conformational B-cell epitopes. CBTOPE software has been developed to predict conformational B-cell epitopes (http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen. html). ElliPro software is available and it is designed for prediction of both linear and conformational B-cell epitopes. I-TASSER server was used to predict 3D structure of protein and then the structure was utilized as input for ElliPro predictions. ElliPro produces a Protrusion Index (PI) score for each predicted epitope.[19]

Prediction of MHC Class I and Class II binding. Cell-mediated immunity plays critical roles in disease resistance and development of vaccine protection against NDV.[20,21] T cell activation is achieved by antigen presentation through major histocompatibility complex (MHC).[22] Therefore, MHC-peptide binding is a necessity for activation of cellular immunity. Chicken MHC is simpler and smaller than that of mammalian and its properties are almost well described.[23] Polymorphic chicken MHC class I and class II genes are present in MHC-B and MHC-Y regions of microchromosome 16 (GGA 16). MHC class I and class II genes are called as B-F and B-L genes in chicken, respectively. Some conserved areas of human leukocyte antigens (*HLA*) alleles and chicken MHC demonstrate similarity and common ancestry.[24] Hence, human MHC was used for MHC binding predictions through bioinformatics tools.[25]

The interaction of T lymphocytes and MHC molecules is known as very main event in the immune responses.[26] Cytotoxic T lymphocytes (CTL) bind to MHC I (HLA A, B and C), and helper T lymphocytes (HTL) interact with MHC II (HLA DR, DP and DQ) alleles.[26] Some tools in IEDB was applied to predict T-cell epitopes which bind to MHC class I and class II molecules.

MHC Class I. To obtain peptide prediction, reference sequence was inserted into IEDB server and artificial neural network (ANN) method was selected (http://tools.iedb.org/bcell/). Information was acquired based on IC50 in 9-mer-length epitopes. IC50 values are grouped into three categories including IC50 value below 50.00 nM determines epitopes with high affinity, between 50 and 500 nM shows intermediate affinity and IC50 > 500 nM defines low affinity epitopes.

MHC Class II. NN-align tool (http://tools.iedb.org/bcell/) was utilized to assess the MHC-II binding prediction using human allele reference set.[24,27-32]

Homology modelling. Three dimensional (3D) structure of fusion protein was obtained by Swiss-model a computational homology modelling server, and then UCSF Chimera visualization tool (version 1.11.2; RBVI, San Francisco, USA) was used to locate and visualize the predicted B- and T-Cell epitopes in structural level.[33]

Molecular docking. MHC class I epitopes were used as ligands. These epitopes were determined to overlap with MHC class II and B-Cell epitopes. Three dimensional model of epitopes was performed by PEP-FOLD server. Two B-F alleles (receptors), *BF2*2101* and *BF2*0401*, were introduced by Koch *et al.* and Zhang *et al.*[29,34] The 3D structures of *BF2*2101* and *BF2*0401* were retrieved from UniProtKB with PDB ID of 3BEW and 4G42, respectively.

Molecular docking is a continuously evolving technique which assists study of peptide-protein interactions. HADDOCK, which adopts a data-driven approach to docking, was performed in this study.[33] Both side chain and backbone conformational changes can be handled in HADDOCK to obtain considerable structural flexibility.[35] Visualization was carried out using the PyMOL molecular graphics system (version 1.8; Schrödinger LLC, New York, USA) and UCSF-Chimera tools. PyMOL was also used to visualize H bonds.

Results

Characterization of F protein. The NDV fusion protein has 553 amino acid residues with a 31-residue signal peptide. This 59.00 kDa protein was membrane bounded and its predicted isoelectric point (pI) was 8.34. InterProScan software determined two cytoplasmic domains (1-116 and 528-553), one non-cytoplasmic domain (143-500) and two trans-membrane domains (117-139 and 503-525) in this protein. PSIPRED calculated composition showed 41.23% α-helix, 19.89% β strand and 38.88% loop region.

Prediction of linear B cell epitopes. Reference sequence was submitted and potentially linear B cell epitopes were obtained. Epitopes with VaxiJen scores above 0.46 were considered as immune-protective epitopes and their conservancy were checked consequently. Seventeen out of 34 BepiPred-predicted epitopes were defined with VaxiJen score. Among those, nine epitopes were antigenic. Three epitopes were in variable region of the protein. Therefore, they were removed from the list of conserved protective epitopes. The other 6 epitopes are tabulated in Table 1. Considering VaxiJen scores, EFDATYQK is the most probable protective epitope predicted by BepiPred. Using BCpreds, 10 linear epitopes were predicted and 4 of them were characterized as being conserved and antigenic. Among those, VSTTKGYASALVPKVVTQVG had the highest score. ElliPro software was served and 8 epitopes were identified in F protein. Antigen probability demonstrated four antigenic epitopes. Two of them were in the conserved regions. All antigenic epitopes identified by BepiPred, BCpreds and ElliPro are shown in Table 1.

Prediction of conformational B cell epitopes. We acquired eleven conformational epitopes based on CBTOPE server, two of which were removed as being located in variable regions (Table 2). Another prediction tool for conformational B cell epitopes is ElliPro. For F protein, 5 discontinuous peptides were obtained by ElliPro. Two of them with PI values of 0.70 or more were chosen and shown in Table 2. The linear B cell epitopes were discovered by the above-mentioned tools and overlapped at one region,[308] VSTTKGY[314], with a VaxiJen score of 0.86. This epitope was also identified as a conformational epitope (ElliPro-predicted).

[376]EGALTT[381] and [445]QKNI[448] were predicted by both CBTOPE and ElliPro as conformational B cell epitopes. [376]EGALTT[381] shared sequences with a linear B cell epitope ([366]NTSACMYSKTEGALTTPYMA[385]).

All epitopes from Table 1 and Table 2 were then analyzed to find any overlapping epitope. Epitope [354]PMSPGIY[360] was recognized as being both BepiPred- and CBTOPE-predicted epitope.

[186]VNDQFNNTARELDCI[200] was also an overlapped linear ElliPro- and CBTOPE-predicted epitope.

[439]EFDATYQK[446] and [376]EGALTTPYMA[385] were discovered to be both linear (BepiPred/BCpreds, respectively) and conformational (ElliPro) epitopes.

Prediction of T cell epitopes. Reference sequence was subjected to IEDB T-cell tools and all MHC class I and II alleles were investigated separately. The output was classified into two lists according to IC50 values (IC50<50 and 50 < IC50 < 500) for each class. As we obtained extremely large amounts of data, it was not possible to include them all, yet some noteworthy data, the overlapped MHC class I and class II epitopes are presented. TAAQITAAA, AQITAAAAL, SPALTQLTI, TQLTIQALY, MRATYLETL, LYCTRIVTF and YSKTEGALT were defined as the overlapped MHC class I and class II T cell epitopes with high binding affinity (IC50 < 50). TQLTIQALY had the highest VaxiJen score and AQITAAAAL had high affinity to interact with five alleles. Twenty-five the overlapped MHC class I and class II T cell

epitopes with intermediate binding affinity (50 < IC50 < 500) were predicted. QQVGVELNL, DCIKITQQV and LTQL TIQAL were of high antigenic potential in this category. YLETLSVST, MRATYLETL, LTQLTIQAL, ITSPALTQL, LYLTELTTV, SIAATNEAV and VELNLYLTE were predicted epitopes which interacted with six alleles or more. MRATYLETL interacted with 11 different MHC class I and II alleles with high and intermediate binding affinity.

Prediction of overlapped T- and B-cell epitopes. There were 13 overlaps between T- and B-cell epitopes including ^{157}SIAATNEAV^{165}, ^{194}TARELDCIK^{201}, ^{199}DCIKITQ QV^{206}, ^{204}QQVGVELNL^{211}, ^{212}LYLTELTTV^{220}, ^{230}LTQLTIQ AL^{238}, ^{298}MRATYLETL^{306}, ^{302}YLETLSVST^{310}, ^{345}LYCTRIVT F^{353}, ^{379}LTTPYMALK^{378}, ^{436}LSGEFDATY^{444}, ^{371}YSKTEGALT^{380}, ^{408}ISQNYGEAV^{416}; among these, ^{212}LYLTELTTV^{220}, ^{194}TARE LDCIK^{201}, ^{199}DCIKITQQV^{206}, ^{204}QQVGVELNL^{211}, ^{436}LSGEFDA TY^{444}, ^{379}LTTPYMALK^{378} and ^{302}YLETLSVST^{310} are linear B cell and T cell epitopes. ^{157}SIAATNEAV^{165}, ^{230}LTQLTIQ AL^{238}, ^{298}MRATYLETL^{306}, ^{345}LYCTRIVTF^{353}, ^{371}YSKTEGAL T^{380}, ^{408}ISQNYGEAV^{416}, ^{436}LSGEFDATY^{444}, ^{379}LTTPYMAL K^{378}, ^{302}YLETLSVST^{310}, ^{194}TARELDCIK^{201} and ^{199}DCIKITQQ V^{206} are conformational B cell and T cell epitopes. ^{436}LSGEFDATY^{444}, ^{379}LTTPYMALK^{378}, ^{194}TARELDCIK^{201}, ^{199}DCIKITQQV^{206} and ^{302} YLETLSVST^{310} were both linear

Table 1. Protective linear B-cell epitopes of F protein identified by BepiPred, BCpreds and ElliPro.

Linear B cell	No.	Position	Epitope sequence	VaxiJen score
BepiPred	1	288-297	NLPSVGNLNN	0.84
	2	307-318	SVSTTKGYASAL	0.54
	3	328-334	SVIEELD	0.64
	4	354-360	PMSPGIY	0.87
	5	439-446	EFDATYQK	1.13
	6	461-477	LDISTELGNVNNSISNA	0.73
BCPREDS	1	180-199	GKMQQFVNDQFNNTARELDC	0.50
	2	214-233	LTELTTVFGPQITSPALTQL	0.45
	3	308-327	VSTTKGYASALVPKVVTQVG	0.54
	4	366-385	NTSACMYSKTEGALTTPYMA	0.46
ElliPro	1	186-211	VNDQFNNTARELDCIKITQQVGVELN	1.04
	2	308-314	VSTTKGY	0.86

Table 2. Conformational B cell epitopes of F protein (CBTOPE, ElliPro).

Conformational B cell	No.	Start	End	Peptide	No. residues	score
CBTOPE	1	72	81	DKEACAKAPL	10	
	2	153	200	RLKESIAAT NEAVHEVTDG LSQLSVAVGK MQQFVNDQFN NTARELDCI	48	
	3	230	231	LT	2	
	4	284	286	GIQ	3	
	5	296	305	NNMRAT YLET	10	
	6	342	381	DLDLYCTRIV TFPMSPGIYS CLSGNTSACM YSKTEGALTT	40	
	7	445	455	QKNISIL DSQV	11	
	8	463	463	I	1	
	9	523	523	C	1	
ElliPro	1			A:N476, A:D479, A:K480, A:A482, A:E483, A:S484, A:N485, A:S486, A:K487, A:L488, A:E489, A:K490, A:V491, A:N492, A:V493, A:R494, A:L495, A:496, A:S497, A:T498, A:S499, A:A500, A:L501, A:I502, A:T503, A:Y504, A:I505, A:V506, A:L507, A:T508, A:V509, A:I510, A:S511, A:L512, A:V513, A:F514, A:G515, A:A516, A:L517, A:S518, A:L519, A:G520, A:L521, A:A522, A:C523, A:Y524, A:L525, A:M526, A:Y527, A:K528, A:Q529, A:K530, A:A531, A:Q532, A:Q533, A:K534, A:T535, A:L536, A:L537, A:W538, A:L539, A:G540, A:N541, A:N542, A:T543, A:L544, A:D545, A:Q546, A:M547, A:R548, A:A549, A:T550, A:T551, A:R552, A:A553	75	0.874
	2			A:A39, A:L306, A:S307, A:V308, A:S309, A:T310, A:T311, A:K312, A:G313, A:Y314, A:E376, A:G377, A:A378, A:L379, A:T380, A:T381, A:P382, A:Y383, A:M384, A:A385, A:L386, A:K387, A:G388, A:S389, A:V390, A:I391, A:A392, A:N393, A:C394, A:K395, A:I396, A:T397, A:T398, A:C399, A:R400, A:C401, A:T402, A:D403, A:P404, A:P405, A:G406, A:I407, A:I408, A:S409, A:Q410, A:N411, A:Y412, A:G413, A:E414, A:A415, A:V416, A:S417, A:L418, A:I419, A:D420, A:R421, A:H422, A:S423, A:C424, A:N425, A:V426, A:L427, A:S428, A:L429, A:D430, A:G431, A:I432, A:T433, A:L434, A:R435, A:L436, A:S437, A:G438, A:E439, A:F440, A:D441, A:A442, A:T443, A:Y444, A:Q445, A:K446, A:N447, A:I448	83	0.70

and conformational B cell epitopes besides being T cell epitopes. [298]MRATYLETL[306] and [345]LYCTRIVTF[353] had IC50 values below 50.

Molecular docking of the best predicted epitopes. All 13 proposed epitopes were introduced as the overlapped CTL, HTL and B-cell epitopes and docked to B-F alleles. Results showed that all peptides exhibited good HADDOCK scores. The score was described as the weighted sum of four energy terms including van der Waals energy, electrostatic energy, distance restraint energy and desolvation energy ($1\ E_{vdw} + 0.2\ E_{elc} + 0.1\ E_{dist} + 1\ E_{solv}$). A lower HADDOCK score represented a better binding condition of peptide-B-F allele complex. The hydrogen bonding interactions of the ligand-receptor complexes were investigated.

Molecular docking of control peptides with *BF2*2101* and *BF2*0401.* The result of control dockings are shown in Table 3. Peptide 'RRKWRRWHL' was selected as positive control based on its lowest HADDOCK score. The proposed epitopes were then docked into the binding groove of the same MHC alleles and the results were compared to those of the positive control.

Table 3. Molecular docking of control peptides with *BF2*2101* and *BF2*0401.*

Epitope	Receptor	No. Cluster	HADDOCK score	H-bond
TPYDINQML	**BF2*2101(3BEW)**	1	−94.30 ± 5.50	10
QYDDAAVYKL		2	105.80 ± 6.50	8
NPRAMQALL		1	80.50 ± 2.60	5
VMAPRTVLL		1	95.00 ± 4.40	5
RIIPRHLQL		2	105.80 ± 6.50	10
GILGFVFTL		1	95.80 ± 1.90	4
EEPTVIKKY		1	93.10 ± 2.60	9
RRKWRRWHL		2	114.50 ± 4.20	4
TPYDINQML	**BF2*0401(4G42)**	1	107.60 ± 7.50	9
QYDDAAVYKL		1	119.40 ± 2.70	10
NPRAMQALL		2	91.60 ± 4.70	7
VMAPRTVLL		1	83.00 ± 1.90	6
RIIPRHLQL		1	97.00 ± 2.90	11
GILGFVFTL		1	104.60 ± 1.10	6
EEPTVIKKY		1	101.00 ± 7.10	10
RRKWRRWHL		1	117.70 ± 2.60	8

Table 4. Molecular docking of epitopes with *BF2*2101, BF2*0401.*

Epitope	Receptor	No. Cluster	HADDACK Score	H-bond
LSGEFDATY	BF2*2101 (3BEW)	1	−126.90 ± 8.00	5
LYLTELTTV		3	96.20 ± 5.600	7
LTTPYMALK		1	107.10 ± 3.60	10
SIAATNEAV		2	88.90 ± 5.90	6
LTQLTIQAL		2	102.20 ± 6.70	10
MRATYLETL		1	125.60 ± 2.90	11
LYCTRIVTF		1	130.60 ± 8.10	8
TARELDCIK		2	95.20 ± 5.60	9
DCIKITQQV		2	103.20 ± 3.9	11
QQVGVELNL		1	101.60 ± 3.90	7
YSKTEGALT		2	87.90 ± 6.50	7
ISQNYGEAV		3	93.80 ± 10.30	15
YLETLSVST		1	86.40 ± 2.20	4
LSGEFDATY	BF2*0401 (4G42)	1	115.10 ± 3.90	9
LYLTELTTV		1	117.80 ± 3.9	10
LTTPPYMALK		1	98.50 ± 5.90	10
SIAATNEAV		1	100.00 ± 1.80	7
LTQLTIQAL		1	96.60 ± 3.30	10
MRATYLETL		1	114.60 ± 5.00	11
LYCTRIVTF		1	124.40 ± 5.50	7
TARELDCIK		1	102.70 ± 3.70	11
DCIKITQQV		2	92.40 ± 2.00	9
QQVGVELNL		1	1020 ± 7.10	13
YSKTEGALT		1	85.40 ± 1.00	9
ISQNYGEAV		1	95.10 ± 0.40	12
YLETLSVST		1	105.60 ± 5.00	14

Molecular docking of epitopes with *BF2*2101.* Best HADDOCK score was generated by LYCTRIVTF- *BF2*2101* complex (Table 4). Eight hydrogen bonds were present between this ligand and receptor (*BF2*2101*). LSGEFDATY and MRATYLETL were ranked high. Other epitopes were scored well (lower than -88). Fifteen hydrogen bonds were formed between ISQNYGEAV and *BF2*2101* (Fig. 1).

Molecular docking of epitopes with *BF2*0401.* LYCTRIVTF- *BF2*0401* showed the highest HADDOCK score of 124.40 ± 5.50 with seven hydrogen bond formations (Table 4). While all other docked peptides were obtained satisfactory scores, LYLTELTTV, LSGEFDATY, and MRATYLETL were demonstrated more favorable docking scores of 117.80 ± 3.90, 115.10 ± 3.90 and 114.60 ± 5.00, respectively (Fig. 1). Detailed docking results are tabulated below.

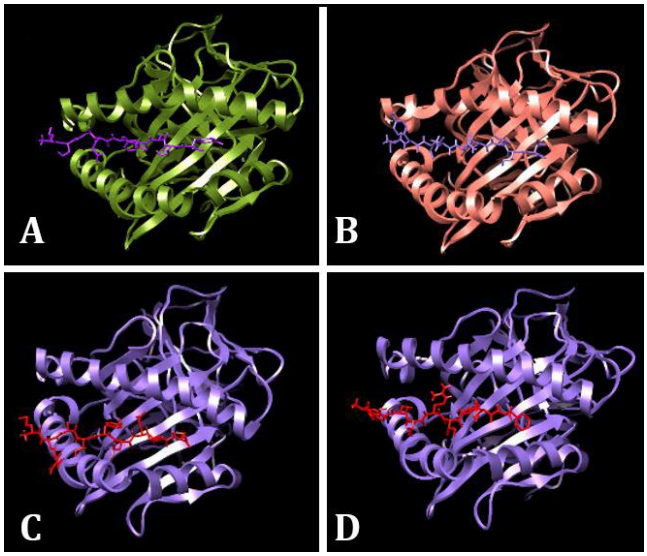

Fig. 1. The predicted epitopes of NDV fusion protein docked to *BF2*0401* and *BF2*2101.* **A)** LSGEFDATY-*BF2*0401* complex, **B)** LYLTELTTV-*BF2*0401* complex, **C)** MRATYLETL-*BF2*2101* complex, and **D)** LYCTRIVTF-*BF2*2101* complex.

Discussion

Besides being efficacious, vaccination is an economically profitable intervention that has reduced burden of infectious diseases significantly during recent decades. Conventional vaccines have been in live-attenuated and inactivated forms. Such vaccines may comprise many proteins whereas immunity is achieved *via* some certain proteins. Thus, additional unnecessary proteins may induce allergic responses in vaccine recipients.[36] In addition, the development of traditional vaccines usually require large budgets. To avoid such limitations, many researchers have focused on development of "peptide vaccines" which contain protective epitopes. Peptide vaccines can be designed to elicit B-cell and T-cell responses and as computational approaches are served to predict candidate epitopes, and development of such vaccines is considered to be safe and low-cost.[35,37,38]

NDV envelope glycoproteins were studied to select the best protein which induces appropriate immune responses. F protein of NDV mediates fusion of viral envelope with cellular membrane along with HN protein. However, some mutations in F protein can increase fusogenic activity of the virus and alter the requisite role of HN protein in promoting fusion.[39] Fusion protein is identified as very important determinant of pathogenicity and its cleavage is responsible for virus virulence. F protein is also the main neutralization antigen and is shown to provide more protective immunity than HN glycoprotein.[40,41]

Active immunity to NDV consists of innate and adaptive immunity. Innate immunity and interferon gamma can induce cell-mediated immunity (CMI) following infection. Some studies have detected CMI responses soon after vaccination. CMI has been regarded as being associated with reduced viral shedding. Another arm of adaptive immune system is humoral immunity which plays a major role in protection against NDV through neutralizing antibodies. Therefore, we discovered both B- and T-cell epitopes as vaccine candidates for NDV fusion protein.[7,28,37,42]

Numerous tools for epitope prediction have been developed over the course of last two decades. Different prediction algorithms were implemented in this study to predict with greater accuracy.[43,44] Superimposition of B-cell epitopes with HTL and CTL epitopes elicits not only good antibody response (due to helper memory immune response) but also proper T cell responses and T cell memory.[45]

Our newly submitted sequence was chosen as reference sequence and we additionally retrieved 126 sequences from all over Asia which provided an insight into conserved regions of genotype VII fusion protein. Linear B-cell prediction was accomplished using three different online softwares. Six, four and two linear B cell epitopes were predicted using BepiPred, BCPREDS and ElliPro, respectively. VSTTKGY was introduced as the overlapped epitope based on the three mentioned tools. VSTTKGY was also discovered to be a conformational B-cell epitope based on ElliPro results. ^{376}EGALTT^{381} was predicted by CBTOPE and ElliPro as a conformational epitope. It was a linear B-cell epitope as well. VSTTKGY and ^{376}EGALTT^{381} shared sequences with T-cell epitopes (YLETLSVST and YSKTEGALT). ^{186}VNDQFNNTARELDCI^{200} and ^{439}EFDATYQK^{446} were determined to be both conformational and linear epitopes.

In a study in 1989, location of neutralizing epitopes of F protein was recognized using monoclonal antibodies (MAbs). Our study demonstrated that five epitopes on F protein (epitopes A1 to A5). Our findings of CBTOPE

algorithm included the same residues. Residues 72, 78, 79, 157-171 and 343 were present in CBTOPE-predicted epitopes at position 72-81 and position 153-200 (Table 2), while a K to R residue substitution was identified at position 78 In sub-genotype VIIg and VIIh.[37] In our study, residues 157-165(SIAATNEAV) was detected as a T-cell epitope. It interacted with both MHC class I and II.

For T-cell epitope investigations, 42 conserved CTL and 103 HTL epitopes were determined. CTL epitopes interacted with MHC class I alleles and HTL epitopes interacted with MHC class II. Chicken MHC is not available in T-cell prediction servers and human MHC alleles are used to find T-cell epitopes instead.[46] The overlapped T-cell epitopes interacting with both human MHC I and II were selected. Such overlapped epitopes enhanced the possibility of antigen presentation to immune system. Finally, seven T-cell epitopes with high binding affinity and 25 T-cell epitopes with intermediate binding affinity were identified. Overlaps with B-cell epitopes were then found.

Chicken MHC and B complex comprised several classes among which B-F and B-L homologous to mammalian MHC class I and class II, respectively. *BF2*2101* (from B21 haplotype) and *BF2*0401* (from B4 haplotype) are defined as chicken MHC class I (*B-F*) alleles.[29,34,47,48]

Taken together, our findings proposed epitopes capable of stimulating B- and T-cell responses using bioinformatics tools and then, with the help of docking simulation, binding of such epitopes to chicken MHC I (B-F) alleles were performed to provide more reliable predictions. 8 docking runs were performed with control peptides. HADDOCK score of our 13 epitopes were relatively in the same range as those of control peptides. One peptide with the lowest HADDOCK score from the collection of control peptides was selected as positive control. Positive control gained a HADDOCK score of 114.50 ± 4.20 and four hydrogen bond formations with *BF2*2101* allele and a HADDOCK score of 117.70 ± 2.60 and eight hydrogen bond interactions with *BF2*0401* allele. Comparing docking results of our epitopes with positive control, we found them of equal or comparatively equal scores.

LYCTRIVTF ranked as the highest based on HADDOCK score (130.60 ± 8.10 with *BF2*2101* and 124.40 ± 5.50 with *BF2*0401*). Docked complex of LYCTRIVTF-*BF2*2101* and LYCTRIVTF-*BF2*0401* showed eight and seven hydrogen bond formations. This epitope was defined as a CTL, HTL and conformational B-cell epitope before molecular docking. LYCTRIVTF was predicted to bind to MHC alleles with high affinity. The IC50 value of <50 nM was used as the threshold for designating high binding epitopes. MRATYLETL was a conformational B- and T-cell epitope which showed interaction with 11 MHC alleles (4 MHCI and 7 MHC II alleles) representing different binding affinities. MRATYLETL-*BF2*2101* and MRATYLETL-*BF2*0401* complexes received docking scores of 125.60 ± 2.90 and 114.60 ± 5.00. LSGEFDATY and LTTPPYMALK were found to be both linear and conformational B-cell epitopes. These two epitopes were bonded to different MHC alleles of class I and II. LSGEFDATY and LTTPPYMALK obtained HADDOCK score of −126.90 ± 8.00 and −107.10 ± 3.60 with *BF2*2101* and 115.10 ± 3.90 and 98.50 ± 5.90 with BF2*0401, respectively. LYLTELTTV and DCIKITQQV were demonstrated to be linear B-cell epitopes (BCpred-predicted and Ellipro-predicted, respectively). They interacted with a variety of MHC alleles. LYLTELTTV was able to bind to 10 MHC alleles. DCIKITQQV presented the highest antigenic value (VaxiJen score: 1.97). Calculated HADDOCK score for LYLTELTTV-*BF2*2101* and LYLTELTTV-*BF2*0401* complexes were 96.20 ± 5.60 and 117.80 ± 3.90, respectively. DCIKITQQV received dock scores of 103.20 ± 3.90 and 92.40 ± 2.00 with the two chicken MHC alleles. SIAATNEAV was introduced as a neutralizing epitope by Yusoff *et al*. CBTope server predicted this epitope as a conformational B-cell epitope. It can elicit T-cell responses as well. Molecular docking of this ligand (SIAATNEAV) into receptors (*BF2*2101* and *BF2*0401*) exhibited proper results.[49] These seven epitopes formed a collection of "high ranked epitopes".

In an attempt to predict candidate epitopes for peptide-vaccine, Badawi *et al.,* conducted a study to identify best B- and T-cell epitopes from fusion protein. YLTELTTVF, NYGEAVSLI, NTSACMVSK and VAVGKMQQF were determined as T-cell epitopes. YLTELTTVF and NYGEAVSLI overlapped with our findings.[13]

To date, computational approaches are immense importance, especially when they come to vaccine design. Although epitope identification has chiefly focused on eliciting humoral responses, lately vaccines based on T-cell epitopes have shown promising results as Khan *et al.,* experimentally validated *in silico*-driven epitopes.[50] Therefore, antigenic B- and T-cell epitopes can be determined by *in silico* analysis and constitute peptide vaccines. We analyzed F protein and finally identified seven epitopes. Although epitope 'LYCTRIVTF' and 'MRATYLETL' showed more promising results, we concluded that the other five epitopes were desirable enough to be experimentally tested. The seven 'high ranked epitopes' need to be tested by subsequent *in vitro* and *in vivo* experiments, so their efficiency as immunogens will be assessed properly.

Acknowledgments

The authors wish to thank Dr. Mohammad Sadegh Saberi for the technical assistance.

Conflict of interest

The authors declare that there are no conflicts of interest related to this article

References

1. Alexander DJ, Bell JG, Alders RG. A technology review: Newcastle disease, with special emphasis on its effect on village chickens. FAO Animal Production and Health Paper No. 161. Rome, Italy:FAO. 2004; 63.
2. Miller PJ, Decanini EL, Afonso CL. Newcastle disease: evolution of genotypes and the related diagnostic challenges. Infect Gent Evol 2010; 10(1):26-35.
3. Aldous EW, Mynn JK, Banks J, et al. A molecular epidemiological study of avian paramyxovirus type 1 (Newcastle disease virus) isolates by phylogenetic analysis of a partial nucleotide sequence of the fusion protein gene. Avian Pathol 2003; 32(3):239-256.
4. Diel DG, da Silva LHA, Liu H, et al. Genetic diversity of avian paramyxovirus type 1: proposal for a unified nomenclature and classification system of Newcastle disease virus genotypes. Infect Genet Evol 2012;12(8):1770-1779.
5. Liu H, Zhao Y, Zheng D, et al. Multiplex RT-PCR for rapid detection and differentiation of class I and class II Newcastle disease viruses. J Virol Methods 2011; 171(1):149-155.
6. Soñora M, Moreno P, Echeverría N, et al. An evolutionary insight into Newcastle disease viruses isolated in Antarctica. Arch Virol 2015;160(8): 1893-1900.
7. Miller P, Koch G et al. Newcastle disease, other paramyxoviruses, and avian metapneumovirus infections. In: Swayne DE, Glisson JR, Mc Dougald LR et al. (Eds) Diseases of poultry 13th ed. Ames, USA: Wiley-Blackwell 2013:138.
8. Dimitrov KM, Ramey AM, Qiu X, et al. Temporal, geographic, and host distribution of avian paramyxovirus 1 (Newcastle disease virus). Infect Genet Evol 2016;39: 22-34.
9. Rui Z, Juan P, Jingliang S, et al. Phylogenetic characterization of Newcastle disease virus isolated in the mainland of China during 2001-2009. Vet Microbiol 2010;141(3-4):246-257.
10. Florea L, Halldorsson B, Kohlbacher O, et al. Epitope prediction algorithms for peptide-based vaccine design. Proc IEEE Comput Soc Bioinform Conf 2003;2:17-26.
11. Dash R, Das R, Junaid M,et al. In silico-based vaccine design against Ebola virus glycoprotein. Adv Appl Bioinform Chem 2017;10:11-28.
12. Osman MM, El Amin EE, Ahmad AA, et al. *In silico* design of epitope based peptide vaccine against virulent strains of HN-Newcastle disease virus (NDV) in poultry species. J Vaccines Vaccin 2016, 7:3 (Suppl). doi:10.4172/2157-7560.C1.040.
13. Badawi MM, Fadl Alla AA, Alam SS, et al. Immunoinformatics predication and in silico modeling of epitope-based peptide vaccine against virulent Newcastle disease viruses. Am J Infect Dis Microbiol 2016; 4(3):61-71.
14. Kim S-H, Wanasen N, Paldurai A, et al. Newcastle disease virus fusion protein is the major contributor to protective immunity of genotype-matched vaccine. PloS One 2013; 8(8):e74022. doi: 10.1371/journal. pone.0074022.
15. McGinnes LW, Reitter JN, Gravel K, et al. Evidence for mixed membrane topology of the Newcastle disease virus fusion protein. J Virol 2003; 77(3):1951-1963.
16. Boroomand Z, Jafari RA, Mayahi M. Molecular characterization and phylogenetic study of the fusion genes of Newcastle disease virus from the recent outbreaks in Ahvaz, Iran. Virusdisease 2016; 27(1):102-105.
17. Ansari HR, Raghava GPS. Identification of conformational B-cell epitopes in an antigen from its primary sequence. Immunome Res 2010; 6:6. doi: 10.1186/1745-7580-6-6.
18. El-Manzalawy Y, Honavar V. Recent advances in B-cell epitope prediction methods. Immunome Res 2010;6 (Suppl 2):S2. doi: 10.1186/1745-7580-6-S2-S2.
19. Ponomarenko J, Bui HH, Li W, et al. ElliPro: a new structure-based tool for the prediction of antibody epitopes. BMC Bioinformatics 2008; 9:514. doi: 10.1186/1471-2105-9-514.
20. Keshavarz M, Mirzaei H, Salemi M, et al. Influenza vaccine: where are we and where do we go? Rev Med Virol 2019;29(1):e2014. doi: 10.1002/rmv.2014.
21. Nami S, Mohammadi R, Vakili M, et al. Fungal vaccines, mechanism of actions and immunology: a comprehensive review. Biomed Pharmacother 2019;109:333-344.
22. Mirzaei HR, Pourghadamyari H, Rahmati M, et al. Gene-knocked out chimeric antigen receptor (CAR) T cells: tuning up for the next generation cancer immunotherapy. Cancer Lett 2018;423:95-104.
23. Kaufman J. The simple chicken major histocompatibility complex: life and death in the face of pathogens and vaccines. Philos Trans R Soc Lond B Biol Sci 2000;355(1400):1077-1084.
24. Chen F, Pan L, Chao W, et al. Character of chicken polymorphic major histocompatibility complex class II alleles of 3 Chinese local breeds. Poult Sci 2012;91(5):1097-1104.
25. Kaufman J. Generalists and specialists: a new view of how MHC class I molecules fight infectious pathogens. Trends Immunol 2018;39(5):367-379.
26. Wucherpfennig KW. The structural interactions between T cell receptors and MHC-peptide complexes place physical limits on self-nonself discrimination. Curr Top Microbiol Immunol 2005;296:19-37.
27. Chicz RM, Urban RG, Gorga JC, et al. Specificity and promiscuity among naturally processed peptides bound to HLA-DR alleles. J Exp Med. 1993;178(1):27-47.

28. Kapczynski DR, Afonso CL, Miller PJ. Immune responses of poultry to Newcastle disease virus. Dev Com Immunol 2013;41(3):447-453.
29. Koch M, Camp S, Collen T, et al. Structures of an MHC class I molecule from B21 chickens illustrate promiscuous peptide binding. Immunity 2007; 27(6):885-899.
30. Wallny HJ, Avila D, Hunt LG, et al. Peptide motifs of the single dominantly expressed class I molecule explain the striking MHC-determined response to Rous sarcoma virus in chickens. Proc Natl Acad Sci USA 2006; 103(5):1434-1439.
31. Miller MM, Taylor Jr RL. Brief review of the chicken major histocompatibility complex: the genes, their distribution on chromosome 16, and their contributions to disease resistance. Poul Sci 2016; 95(2):375-392.
32. Chappell P, Meziane El K, Harrison M, et al. Expression levels of MHC class I molecules are inversely correlated with promiscuity of peptide binding. ELife 2015; 4:e05345. doi: 10.7554/eLife.05345
33. Pourseif MM, Moghaddam G, Daghighkia H, et al. A novel B- and helper T-cell epitopes-based prophylactic vaccine against Echinococcus granulosus. Biompacts 2018;8(1):39-52.
34. Zhang J, Chen Y, Qi J, et al. Narrow groove and restricted anchors of MHC class I molecule BF2*0401 plus peptide transporter restriction can explain disease susceptibility of B4 chickens. J Immunol 2012; 189(9):4478-4487.
35. Andre FE, Booy R, Bock HL, et al. Vaccination greatly reduces disease, disability, death and inequity worldwide. Bull World Health Organ 2008; 86(2):140-146.
36. Dreskin SC, Halsey NA, Kelso JM, et al. International consensus (ICON): allergic reactions to vaccines. World Allergy Organ J 2016; 9(1):32. doi: 10.1186/s40413-016-0120-5.
37. Li W, Joshi MD, Singhania S, et al. Peptide vaccine: progress and challenges. Vaccines (Basel) 2014; 2(3):515-536.
38. Langeveld J, Casal JP, Osterhaus AD, et al. First peptide vaccine providing protection against viral infection in the target animal: studies of canine parvovirus in dogs. J Virol 1994;68(7):4506-4513.
39. Johnson JB, Schmitt AP, Parks GD. Point mutations in the paramyxovirus F protein that enhance fusion activity shift the mechanism of complement-mediated virus neutralization. J Virol 2013;87(16):9250-9259.
40. Ganar K, Das M, Sinha S, et al. Newcastle disease virus: current status and our understanding. Virus Res 2014;184:71-81.
41. Ji Y, Liu T, Jia Y, et al. Two single mutations in the fusion protein of Newcastle disease virus confer hemagglutinin-neuraminidase independent fusion promotion and attenuate the pathogenicity in chickens. Virology 2017;509:146-151.
42. Purcell AW, McCluskey J, Rossjohn J. More than one reason to rethink the use of peptides in vaccine design. Nat Rev Drug Discov 2007;6(5):404-414.
43. Lundegaard C, Lund O, Kesmir C, et al. Modeling the adaptive immune system: predictions and simulations. Bioinformatics 2007;23(24):3265-3275.
44. Trost B, Bickis M, Kusalik A. Strength in numbers: achieving greater accuracy in MHC-I binding prediction by combining the results from multiple prediction tools. Immunome Res 2007; 3:5. doi: 10.1186/1745-7580-3-5.
45. Parida R, Shaila MS, Mukherjee S, et al. Computational analysis of proteome of H5N1 avian influenza virus to define T cell epitopes with vaccine potential. Vaccine. 2007;25(43):7530-7539.
46. Hou Y, Guo Y, Wu C, et al. Prediction and identification of T cell epitopes in the H5N1 influenza virus nucleoprotein in chicken. PloS One 2012;7(6):e39344. doi: 10.1371/journal.pone.0039344.
47. de Vries SJ, van Dijk M, Bonvin AMJJ. The HADDOCK web server for data-driven biomolecular docking. Nat Protoc 2010;5(5):883-897.
48. Baelmans R, Parmentier HK, Nieuwland MGB, et al. Serological screening for MHC (B)-polymorphism in indigenous chickens. Trop Anim Health Prod 2005;37:93-102.
49. Yusoff K, Nesbit M, McCartney H, et al. Location of neutralizing epitopes on the fusion protein of Newcastle disease virus strain Beaudette C. J Gen Virol 1989;70 (Pt 11):3105-3109.
50. Khan MK, Zaman S, Chakraborty S, et al. In silico predicted mycobacterial epitope elicits in vitro T-cell responses. Mol Immunol 2014;61(1):16-22.