ISSN No:-2456-2165

Prediction of Conserved Epitopes of Hemagglutinin Neuraminidase from New Castle Disease Virus Isolated in Madagascar between 1992 and 2011: An in Silico Study

Randriamamisolonirina Tendrinarisoa*[1, 2]; Ravoavy Hajarivelo [2]; Andriantsimahavandy Abel [2]; Maminiaina Olivier Fridolin [1, 2, 3]
[1]Department of Research and Quality Controls, Malagasy Institute of Veterinary Vaccines (IMVAVET), Ministry of National Education and Scientific Research, Antananarivo, Madagascar
[2]Department of Applied Biochemistry, Faculty of Sciences, University of Antananarivo, Ministry of National Education and Scientific Research, Antananarivo, Madagascar
[3]Department of Education of Veterinary Science and Medicine, Faculty of Medicine, University of Antananarivo, Ministry of National Education and Scientific Research, Antananarivo, Madagascar

Abstract:- Newcastle disease (ND) is one of the avian viruses that cause considerable economic losses worldwide. It is caused by a virus called Newcastle disease virus (NDV). Vaccines were available to reduce the devastating effect of ND. However, outbreaks of the disease were still found, even in vaccinated flocks. In order to generate a new multi-epitopic vaccine, this study aims to predict all conserved epitopes in the hemagglutinin neuraminidase (HN) of NDV isolated in Madagascar between 1992 and 2011; using *bioinformatics* tools. A total of 12 HN protein sequences were extracted from the National Center of Biotechnology Information (NCBI) database. To know the conserved regions, they were aligned in BIOEDIT 7.2.5. Then, immuno-informatics servers were used to predict the B and T-cells epitopes present in all sequences. The predicted T-cells epitopes are further docked with two chickens Major Histocompatibility Complex (MHC) to find out their interaction capacity. As results, we identified 39 conserved epitopes including 17 B-cells epitopes and 22 T-cells epitopes. The majority of these epitopes are localized at the globular head of HN, forming 11 antigenic sites. Some of them are already identified by other researchers while some others are newly identified in this study. Nevertheless, all of these epitopes could be used in the design of a new vaccine against ND. Since this is a preliminary and *in silico* study, further study is needed to valorize these results.

Keywords:- *Newcastle disease; NDV; Hemagglutinin Neuraminidase; epitopes; Bioinformatics; immuno-informatics.*

I. INTRODUCTION

Newcastle disease (ND) is an infectious avian disease caused by a virus commonly known as Newcastle Disease Virus (NDV) or Avian Paramyxovirus type I (AMPV-1). These viruses belong to the family of Paramyxoviridae and to the *Orthoavulavirus* genus [1,2]. ND infects several species of poultry. However, chickens have been reported to be the most susceptible [3,4]. Due to this situation, it causes a very important economic loss in the poultry industry [5-7]. In Madagascar, this disease is responsible for 49% of mortality in domestic poultry [8,9].

Vaccination remains the best means of protection against ND. In general, the vaccines currently available are manufactured with strains of NDV belonging to genotypes I to III [10,11]. They are live attenuated or killed vaccines and are part of the first generation of vaccines [12]. Some studies have shown that these vaccines have the ability to protect poultry from clinical disease, but vaccinated birds may still shed virulent viruses that could infect healthy non-vaccinated flocks [13-15]. According to Miller et colleagues, this failure is due to the genetic diversity between the strains used in the design and the strains circulating in the field [16,17].

In Madagascar, the circulating strains responsible for the epizootic are classified as genotypes XI [18]. Thus, a significant genetic distance has been observed between these strains and the classical vaccine strains. Consequently, ND outbreaks have always been reported in all regions of the island [19]. Even if the vaccination rate is still lower [20] and that there is a rapid evolution of viruses belonging to the Paramyxoviridae family [21,22], these delicate situations allowed us to consider a reduction in the effectiveness of ND vaccines used in the field. Thus, the generation of a new vaccine, more adapted to the current and future epidemiological appears unavoidable if we want to eradicate this disease.

Following the outbreak of the COVID-19, the discovery of multi-epitope vaccines became popular [23-25]. The common concept of this type of vaccine is based on the approach of combining immuno-dominant B-cell and T-cell epitopes capable of producing specific immunity [26,27]. The

identification of epitopes in the target antigens is therefore one of the crucial steps in the design of this type of vaccine. In the context of ND, previous research was done to identify the immuno-dominant epitopes present in viral proteins, such as fusion protein (F) [28], hemagglutinin neuraminidase (HN) [29], matrix protein (M) [30] and nucleoprotein (N) [31]. In these works, the authors used classical methods, based on monoclonal antibody assays (mAbs). However, these methods are time consuming and very expensive [32].

Recently, bioinformatics and immuno-informatics are the most solicited methods for the research of epitopes vaccinal [33,34]. These methods have also been used in the theoretical design of a recombinant or multi-epitope vaccine [35-37]. In this study, we used these techniques to identify conserved epitopes in NDV strains isolated in Madagascar between 1992 and 2011. Among the structural proteins of NDV, HN was chosen as a target because it can carry the major antigens of the virus and can also induce neutralizing antibodies during infection [38].

II. MATERIALS AND METHODS

A. Sequence retrieval and determination of conserved regions

12 sequences of HN protein from NDV isolated in Madagascar were downloaded from the National Center for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov/protein). Information about these sequences is summarized in Table 1. To determine the conserved regions, the downloaded sequences were submitted to a multiple sequence alignment in BIOEDIT 7.2.5 software, using CLUSTALW program [39]. The sequence of MG-1992 is considered as reference.

N°	Name	Acc. Num	Coll. Date	Gen.	Ref.
1	MG-1992	ADQ64396	1992	XI	[18]
2	MG-725	ADQ64390	2008	XI	[18]
3	MG-MEOLA	ADQ64399	2008	XI	[18]
4	MG-39	ADQ64402	2008	XI	[18]
5	MGMNJ	AGW43230	2009	XI	[40]
6	MGF003C	AGW43224	2010	XI	[40]
7	MGF082T	AGW43226	2010	XI	[40]
8	MGF015C	AGW43225	2011	XI	[40]
9	MGF120T	AGW43227	2011	XI	[40]
10	MGF166	AGW43228	2011	XI	[40]
11	MGF192C	AGW43229	2011	XI	[40]
12	MGS1130T	AGW43231	2011	XI	[40]

TABLE 1: LIST OF STRAINS AND THEIR ACCESSION NUMBERS IN THE NCBI DATABASE

B. Prediction of B-cells epitopes

B-cell epitopes were divided into linear and discontinuous [41]. Linear B-cell epitopes were predicted using two online servers ABCpred [42] and BEPIPRED 2.0 [43]. In ABCpred, we set the amino acid length to 10-mers and the scoring threshold to 0.51. In contrast, the length of amino acid is not fixed in BEPIPRED-2.0 while the default threshold 0.5 was chosen during the prediction. The epitopes predicted by these two servers were assembled into the list of linear B-cell epitopes. In addition, discontinuous B-cell epitopes were predicted in the DISCOTOPE 2.0 server, with a threshold value set at -3.7 [44]. Peptides formed with a minimum of six amino acids residues were chosen as discontinuous B-cell epitopes.

C. T-cells epitopes prediction

T-cells epitopes were predicted using prediction tools in the Immune Epitope Database (IEDB) [45]. Human Major Histocompatibility Complex (MHC) or HLA alleles were used in place of chicken MHC alleles due to the absence of these alleles in IEDB.

For $TCD8^+$ cell epitopes, prediction was performed using the IEDB MHC-I tool (http://tools.iedb.org/mhci/). The binding affinity of peptides to MHC-I molecules was evaluated using the NetMHCpan-4.1 method [46]. The length of the peptides was set to 9-mer because this length is preferable for binding the majority of ligands presented by the HLA alleles [47]. Peptides that interact with more than two MHC-I alleles and have a score ≤ 1% were selected as $TCD8^+$ cells epitopes [48].

For $TCD4^+$ cells epitopes**,** prediction was done using the IEDB MHC-II tool (http://tools.iedb.org/mhcii/). The consensus method, which combined the three methods of prediction including SMM, N.N.align and Combi.lib was used in the analysis [49]. The length of the peptides was set to 15-mers. Peptides that bind to multiple HLA alleles and with a percentil rank less that 10% were selected as $TCD4^+$ cells epitopes [50].

D. Prediction of antigenicity and membrane topology of predicted B and T-cells epitopes

The antigenicity of predicted B and T-cells epitopes was analyzed in the VaxiJen v2.0 server (http://www.ddgpharmfac.net/vaxijen/VaxiJen/VaxiJen.html) using threshold score of 0.4 [51]. Then, sequences with a VaxiJen v2.0 score higher than 0.4 were also submitted to the TMHMM v2.0 server (http://www.cbs.dtu.dk/services/TMHMM/) to check their membrane topologies [52]. In these conditions, peptides with a VaxiJen v2.0 score higher than 0.4 and an exo-membrane topology were retained.

E. Molecular docking

The binding affinity of all previously retained T-cell epitopes to the chicken MHC was evaluated using a molecular docking program. First, two chicken MHC molecules (*BF2*21:01*; *pdb ID = 3BEW* and *BL2*019: 01*; *pdb ID = 6KVM*) were downloaded into the RCSB PDB server (https://www.rcsb.org/). Then, they were treated in Discovery studio software version 2017 (D.S v2017) to get rid co-crystallized peptides and unnecessary water [53]. After

ISSN No:-2456-2165

that, the linear sequences of T cells epitopes were folded in PEP-FOLD 3.5 server (http://bioserv.rpbs.univ-paris-diderot.fr/services/PEP-FOLD3) to obtain their three-dimensional (3D) forms [54]. The sOPEP energy scores were utilized in the selection of the 3D models proposed by PEP-FOLD 3.5 [55]. The 3D structures of the receptors (chicken MHC) and ligands (modeled T-cells epitopes) are then minimized in SWISS-pdb Viewer v4 [39]. Finally, molecular docking was performed using Autodock Vina program [56] *included in AUTODOCK TOOLS* [57]. The results were retrieved as pose and binding energy.

F. Modeling of the 3D structure of HN and vizualization of the position of predicted epitopes

The 3D structure of HN was modeled using the online server SWISS-MODEL (https://swissmodel.expasy.org/interactive). The modeling is based on the homology method [58]. The crystal structure of HN from the Australia Victoria strain (*pdb ID = 3t1e*) was chosen as a template [59]. The obtained 3D model was minimized in the software *SWISS-pdb Viewer v4* [39]. Finally, its stability was measured by Ramachandran profile analysis in D.S v2017 [60,61]. The positions of all predicted B and T cells epitopes were localized in this structural model.

III. RESULTS

A. Residues and conserved domains in HN

The result of the multiple sequences alignment is presented in Figure 1. Considering MG-1992 as a reference and ignoring the amino acid similarity and identity, 40 mutations were found in all aligned sequences. The majority of them were located in the amino-terminal region (amino acids 1-80) and the others were dispersed along the linear sequence. However, the thirteen cysteine residues that form the disulfide bonds and the six asparagine residues that form the six glycosylation sites were all conserved [62-64]. In addition, amino acids located in the fourteen regions 81-116, 121-141, 143-207, 209-219, 221-253, 255-302, 304-314, 326-346, 348-361, 363-376, 391 -436, 444-451, 454-494, and 521--570 were also conserved.

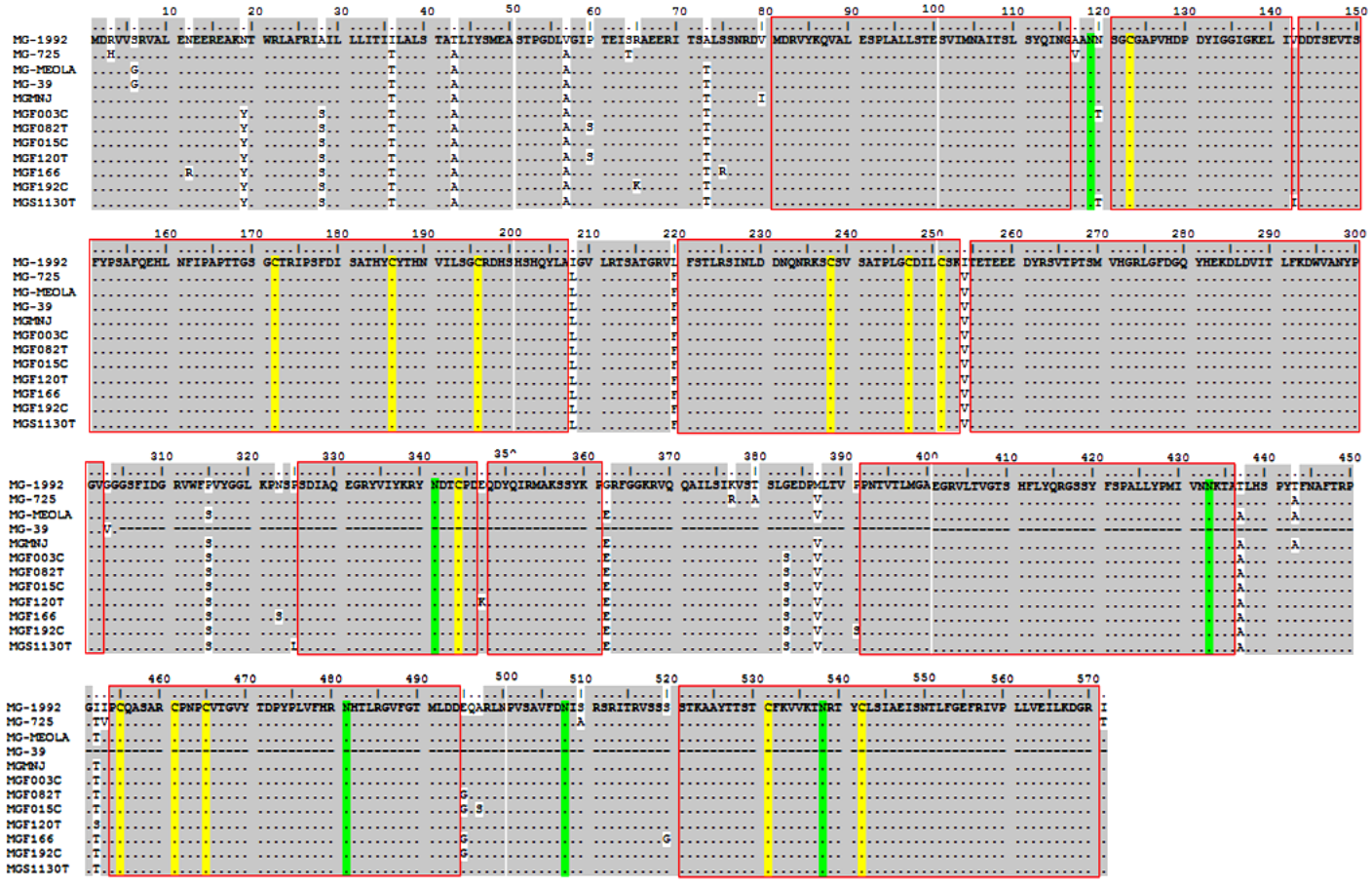


Fig. 1: Residues and conserved regions in the HN of NDV isolated in Madagascar.

The names of the 12 sequences are indicated on the left. MG-1992 is considered as a reference sequence. The conserved amino acids are represented by the dots and are colored in gray. They are represented by a single letter if they are modified. Black dashes indicate gaps in the sequence. The 13 conserved cysteines are highlighted in yellow. In contrast, the six conserved asparagine residues are colored in green. Finally, the 14 conserved regions are framed in red.

ISSN No:-2456-2165

B. Prediction of linear and discontinuous B-cell epitopes

In total, ABCpred and BEPIPRED 2.0 predicted 157 peptides likely to be linear B-cell epitopes. Of these peptides, 69 peptides were located in the conserved regions described in Figure 1. In these 69 conserved peptides, only 46 obtained a VaxiJen score higher than 0.4 (Supplementary Table 1). According to the results of prediction of the membrane topology in the TMHMM v2.0 server, 14 peptides among these 47 were located in exo-membrane regions (Table 2). Therefore, they are retained in the list of linear B cells epitopes.

In addition, 34 peptides were predicted by the DISCOTOPE 2.0 server. Only 14 of these peptides were found in the conserved regions of HN (Supplementary Table 2). Among them, 8 peptides were formed with amino acid (aa) residues smaller than six. Therefore, they were removed from the list of B-cell epitopes. However, the 6 peptides were retained and their antigenicity was analyzed in VaxiJen v2.0 server. As results, they all obtained VaxiJen scores higher than 0.4. However, the prediction of the membrane topology by TMHMM v2.0 server indicated that only 3 peptides among them were located in the exo-membrane regions of the HN protein. Thus, these 3 peptides were retained in the list of conserved discontinuous B-cell epitopes in the 12 sequences of HN analyzed (Table 3).

N°	Start	End	Sequence	Length	VaxiJen v2.0 score	TMHMM v2.0 result
1	92	101	SPLALLSTES	10	0.8382	outside
2	124	133	GAPVHDPDYI	10	0.4296	outside
3	155	164	AFQEHLNFIP	10	1.0412	outside
4	158	167	EHLNFIPAPT	10	1.3725	outside
5	161	170	NFIPAPTTGS	10	0.7259	outside
6	273	282	GRLGFDGQYH	10	1.6473	outside
7	274	283	RLGFDGQYHE	10	1.1843	outside
8	277	286	FDGQYHEKDL	10	1.1298	outside
9	278	287	DGQYHEKDLD	10	1.4379	outside
10	279	288	GQYHEKDLDV	10	1.5970	outside
11	300	309	PGVGGGSFID	10	0.7099	outside
13	318	327	GGLKPNSPSD	10	0.4094	outside
14	447	457	FTRPGIIPCQA	10	0.5708	outside

TABLE 2**:** LIST OF LINEAR B-CELL EPITOPES PREDICTED IN THIS STUDY

N°	Start	End	Séquence	Length	VaxiJen v2.0 score	TMHMM v2.0 result
1	321	332	KPNSP*DIA*EG	12	0.9504	outside
2	322	333	PNSP**IA*EGR	12	0.6735	outside
3	322	332	PNSP*DIA*EG	11	0.8039	outside

TABLE 3 : LIST OF DISCONTINUOUS B- CELL EPITOPES PRESENT IN THE CONSERVED REGIONS OF THE HN PROTEIN

C. Prediction of TCD8$^+$ and TCD4$^+$ cells epitopes

Firstly, IEDB MHC-I tool predicted 109 peptide sequences having a score ≤1 and interacting with five types of HLA alleles. Among them, only 37 were located in the 14 conserved regions of HN (Supplementary Table 3). Secondly, antigenicity analysis in VaxiJen v2.0 server showed that 26 of these 37 conserved peptides had a VaxiJen v2.0 score higher than 0.4. In which, 14 had exo-membrane topology and 12 were localized in the intra-membrane region. Indeed, 14 peptides were retained in the list of TCD8$^+$ cells epitopes of HN protein (Table 4).

Secondly, IEDB MHC-II tool predicted 90 peptides having a score ≤ 10% and interacting with two types of HLA alleles. Of these 90 selected peptides, only 35 peptides are located in the conserved regions of the HN protein (Supplementary Table 4). Of these 35 retained peptides, 11 had a VaxiJen v2.0 score less than 0.4 and 10 were located in an intra-membranous region. Consequently, 21 peptides were eliminated and the remaining 14 (those with a VaxiJen score higher than 0.4 and exo-membrane topology) were retained in the list of epitopes for TCD4$^+$ cells (Table 5).

ISSN No:-2456-2165

N°	Start	End	Peptide	Length	HLA Alleles	Percentile Rank	VaxiJen 2.0 score	TMHMM v2.0 Result
1	148	156	VTSFYPSAF	9	HLA-B*35:01 HLA-B*57:01	0.95 0.72	0.4192	outside
2	154	162	SAFQEHLNF	9	HLA-B*35:01 HLA-B*57:01	0.08 0.27	0.6244	outside
3	177	185	SFDISATHY	9	HLA-B*35:01	0.57	1.7398	outside
4	273	281	GRLGFDGQY	9	HLA-B*27:05	0.04	1.9611	outside
5	299	307	YPGVGGGSF	9	HLA-B*07:02 HLA-B*35:01	0.1 0.04	1.0615	outside
6	406	414	TVGTSHFLY	9	HLA-B*35:01	0.58	0.5265	outside
7	419	427	SYFSPALLY	9	HLA-B*35:01 HLA-B*27:05	0.56	0.6322	outside
8	468	476	GVYTDPYPL	9	HLA-A*02:01	0.54	0.6599	outside
9	470	478	YTDPYPLVF	9	HLA-A*02:01 HLA-B*07:02 HLA-B*35:01 HLA-B*57:01	0.56 0.92 0.08 0.2	1.0589	outside
10	474	482	YPLVFHRNH	9	HLA-B*35:01	0.77	1.1189	outside
11	498	506	RLNPVSAVF	9	HLA-A*02:01 HLA-B*07:02 HLA-B*57:01 HLA-B*35:01	0.91 0.76 0.53 0.56	1.4846	outside
12	545	553	IAEISNTLF	9	HLA-B*35:01	0.53	0.4915	outside
13	551	559	TLFGEFRIV	9	HLA-A*02:01	0.19	0.4567	outside
14	557	565	RIVPLLVEI	9	HLA-A*02:01	0.16	1.1852	outside

TABLE 4: PREDICTED $TCD8^+$ CELLS EPITOPES

N°	Start	End	Length	Peptide	HLA Alleles	Percentile Rank	VaxiJen 2.0 score	TMHMM result
1	86	100	15	KQVALESPLALLSTE	HLA-DRB1*03:01	8.70	0.4735	outside
2	221	235	15	FSTLRSINLDDNQNR	HLA-DRB1*03:01	7.80	0.6869	outside
3	278	292	15	DGQYHEKDLDVITLF	HLA-DRB3*01:01	5.70	0.7537	outside
4	279	293	15	GQYHEKDLDVITLFK	HLA-DRB3*01:01	5.70	0.7371	outside
5	376	390	15	IKVSTSLSEDPVLTV	HLA-DRB3*01:01	6.50	0.4333	outside
6	377	391	15	KVSTSLSEDPVLTVS	HLA-DRB3*01:01	6.60	0.4406	outside
7	378	392	15	VSTSLSEDPVLTVSP	HLA-DRB3*01:01 HLA-DRB1*03:01	7.30 8.50	0.6321	outside
8	379	393	15	STSLSEDPVLTVSPN	HLA-DRB3*01:01 HLA-DRB1*03:01	7.90 8.50	0.6043	outside
9	463	477	15	NPCVTGVYTDPYPLV	HLA-DRB3*01:01	2.80	0.8689	outside
10	464	478	15	PCVTGVYTDPYPLVF	HLA-DRB3*01:01	2.70	1.0015	outside
11	465	479	15	CVTGVYTDPYPLVFH	HLA-DRB3*01:01	2.70	0.7999	outside
12	466	480	15	VTGVYTDPYPLVFHR	HLA-DRB3*01:01	2.70	0.7708	outside
13	467	481	15	TGVYTDPYPLVFHRN	HLA-DRB3*01:01	2.70	0.7173	outside
14	468	482	15	GVYTDPYPLVFHRNH	HLA-DRB3*01:01	9.00	0.8078	outside

TABLE 5: PREDICTED $TCD4^+$ EPITOPES

ISSN No:-2456-2165

D. Molecular docking

The 28 selected T-cell epitopes are docked with two alleles *BF2*21:01* and *BL2*019:01* of MHC classes I and II in chickens. In this case, the 14 $TCD8^+$ cells epitopes are docked to the *BF2*21:01* allele while the 14 $TCD4^+$ cells epitopes are docked to the *BL2*019:01* allele. After docking, the epitope T-MHC complexes with the lowest interaction energy were selected and the binding modes of the epitopes in the peptide cavity of the MHC molecule were visualized in D.S v2017. As results, we found that all docked epitopes T are well anchored in the peptide groove of the chicken MHC molecule (Fig. 2).

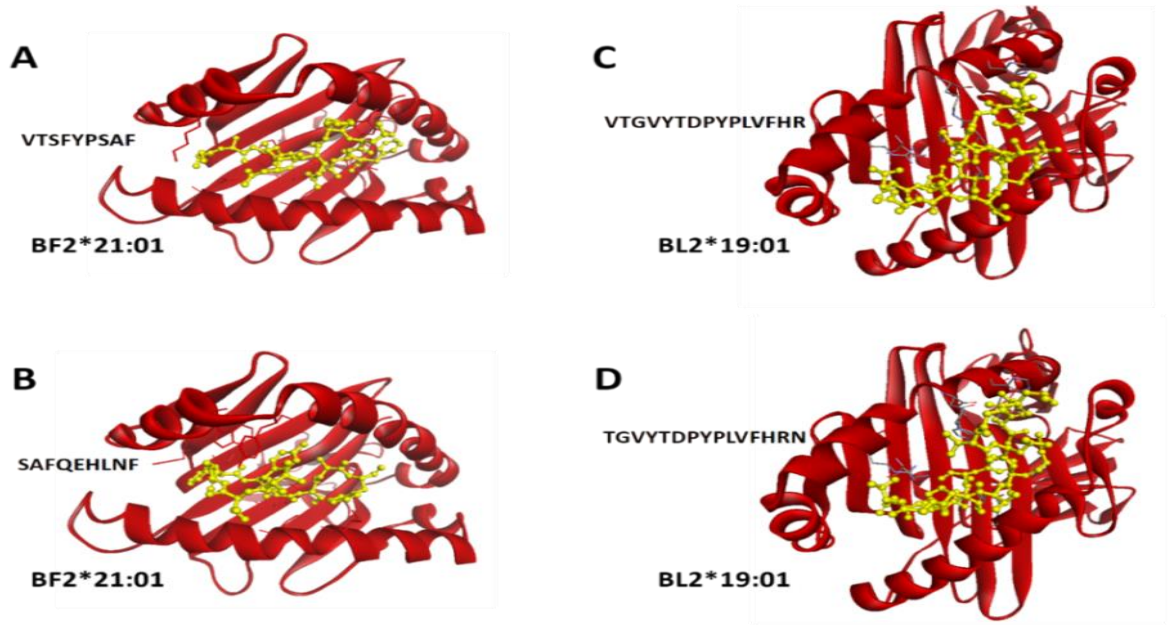


Fig. 2: Mode of binding for T-cell epitopes in chicken MHC peptides cavities: the complexes are visualized in D.S v2017, the receptors (BF2*21:01 and BL2*019:01) are represented as red ribbons while the $TCD4^+$ or $TCD8^+$ epitopes are represented as yellow balls-sticks. (A): VTSFYPSAF - BF2*21:01, (B): SAFQEHLNF- BF2*21:01; (C): VTGVYTDPYPLVFHR- BL2*019:01 and (D): TGVYTDPYPLVFHRN- BL2*019:01

In addition, the binding energy of the epitope T-MHC complexes is noted (Table 6 and 7). It is clear that only 9 $TCD8^+$ cells epitopes and 13 $TCD4^+$ cells epitopes had lower interaction energy than the reference complexes. The rest had higher interaction energy. The 7 peptides with higher binding energy than the reference complexes were eliminated. Thus, 21 peptides were retained from the list of conserved epitopes T of HN in NDV isolated in Madagascar.

Epitope	Start	End	Peptide	Energy of interaction Kcal/mol
Référence	-	-	**REVDEQLLSV**	**-8.2**
1	148	156	VTSFYPSAF	-9.0
2	154	162	SAFQEHLNF	-8.3
3	177	185	SFDISATHY	-8.0
4	273	281	GRLGFDGQY	-8.6
5	299	307	YPGVGGGSF	-8.3
6	406	414	TVGTSHFLY	-9.1
7	419	427	SYFSPALLY	-8,7
8	468	476	GVYTDPYPL	-8.8
9	470	478	YTDPYPLVF	-8.6
10	474	482	YPLVFHRNH	-8.1
11	498	506	RLNPVSAVF	-7.9
12	545	554	IAEISNTLF	-8.1
13	551	559	TLFGEFRIV	-8.1
14	557	565	RIVPLLVEI	-8,1

TABLE 6 : INTERACTION ENERGY OF $TCD8^+$ CELL EPITOPES WITH THE CHICKEN MHC CLASS I ALLELE BF2*21:01

ISSN No:-2456-2165

Epitope	Start	End	Peptide	Energy of interaction Kcal/mol
Reference	-	-	**PGDSDIIRSMPEQTSEK**	**-4.8**
1	86	100	KQVALESPLALLSTE	-2.7
2	221	235	FSTLRSINLDDNQNR	-8.1
3	278	292	DGQYHEKDLDVITLF	-8.2
4	279	293	GQYHEKDLDVITLFK	-7.3
5	376	390	IKVSTSLSEDPVLTV	-8.2
6	377	391	KVSTSLSEDPVLTVS	-7.5
7	378	392	VSTSLSEDPVLTVSP	-7,3
8	379	393	STSLSEDPVLTVSPN	-7.5
9	463	477	NPCVTGVYTDPYPLV	-8.2
10	464	478	PCVTGVYTDPYPLVF	-7.5
11	465	479	CVTGVYTDPYPLVFH	-7.7
12	466	480	VTGVYTDPYPLVFHR	-8.5
13	467	481	TGVYTDPYPLVFHRN	-8.7
14	468	482	GVYTDPYPLVFHRNH	-5.6

TABLE 7 : INTERACTION ENERGY OF TCD4$^+$ CELL EPITOPES WITH THE CHICKEN MHC CLASS II ALLELE BL2*019:01

E. Identification of overlapping epitopes

After all prediction steps, 39 peptides were retained from the list of conserved T and B-cell epitopes of NDV HN protein isolated in Madagascar (Table 2, 3, 6 and 7). Among these, we found 20 overlapping epitopes including 6 B-cell epitopes, 6 TCD8$^+$ cell epitopes and 8 TDC4$^+$ cell epitopes. These overlaps were observed between B/TCD8$^+$ cells epitopes, then between B/TCD4$^+$ cells epitopes and finally between TCD8$^+$/ TCD4$^+$ cells epitopes (Table 8).

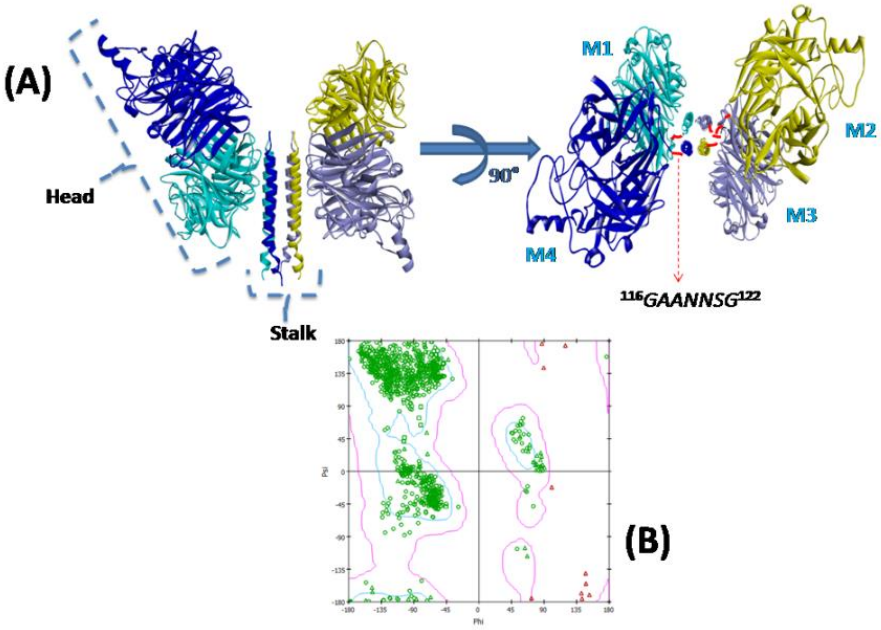
N°	Position		Sequences		
	Start	*End*	*Linear B-cells epitopes*	*TCD8$^+$ cells epitope*	*TCD4$^+$ cells epitope*
1	154	164	^{155} **AFQEHLNFIP** ^{164}	^{148} VTSFYPS**AF** ^{156} ^{154} **SAFQEHLNF** ^{162}	-
2	273	283	^{273} **GRLGFDGQY** ^{282} ^{274} **RLGFDGQY**H ^{283}	^{273} **GRLGFDGQY** ^{282}	-
3	299	309	^{300} **PGVGGGSF**ID ^{309}	^{299} Y**PGVGGGSF** ^{307}	
3	278	293	^{278} **DGQYHEKDLD** ^{287} ^{279} **GQYHEKDLD** ^{288}	-	^{278} **DGQYHEKDLD**VITLF ^{292} ^{279} **GQYHEKDLD**VITLFK ^{293}
5	463	482	-	^{468} **GVYTDPYPL** ^{476} ^{470} **YTDPYPLVF** ^{478}	^{463} NPCVT**GVYTDPYPL**V ^{477} ^{464} PCVT**GVYTDPYPLVF** ^{478} ^{465} CVT**GVYTDPYPLVF**H ^{473} ^{466} VT**GVYTDPYPLVF**HR ^{480} ^{467} T**GVYTDPYPLVF**HRN ^{481} ^{468} **GVYTDPYPLVF**HRNH ^{482}

TABLE 8 : LISTS OF OVERLAPPING EPITOPES PREDICTED IN HN OF NDV ISOLATED IN MADAGASCAR

F. Localization of predicted epitopes in the 3D structure of HN

To localize these 39 predicted epitopes in the 3D structure of the HN protein, the structure was firstly modeled in the SwissModel server. The resulting model was then visualized in the D.S v2017 software. After this visualization, we found that the obtained 3D model was similar to the 3D structure of the HN from the Australia Victoria strain (*pdb ID= 3t1e*). It contained a globular head and a spiral rolled stalk. The assembly is constituted by four monomers denoted M1 to M4 (Fig.3A). The peptide fragments (^{116}*GAANNSG*^{122}) that connect the head and the stalk were not modeled because they were not found in the 3t1e structure model used for modeling. Nevertheless, we found that the resulting 3D structure was stable: no amino acid residues (except for glycines) were found in the unauthorized regions of the Ramachandran plot (Fig.3 B).

ISSN No:-2456-2165



.

Fig. 3: Model 3D of HN protein structure of strain MG-1992 isolated in Madagascar: (A) Model 3D of HN is represented as a solid ribbon. It is composed of a globular head (residues at position 1 -143) and a spirally rolled stalk (residues at position 125-571). Each monomer that forms the structure has been represented with different colors. (B) Ramachandran plot demonstrating the validity of the model structure obtained: no amino acid residues were found in the unauthorized regions of Ramachandran plot, except for glycines (red triangles)

The visualization of the location of the 39 epitopes predicted in the structural model of HN revealed that they are all placed at the globular head of the protein, with the exception of the linear B-cell epitope [92] *SPLALLSTES* [101] which is located at the stalk (Fig. 4A and B). Furthermore, some overlap was also found on the placements of these epitopes. Accordingly, the 39 epitopes predicted in this study were formed 11 antigenic sites including 3 sites for overlapping epitopes and 8 sites for non-overlapping epitopes (Fig. 4A and 4B).

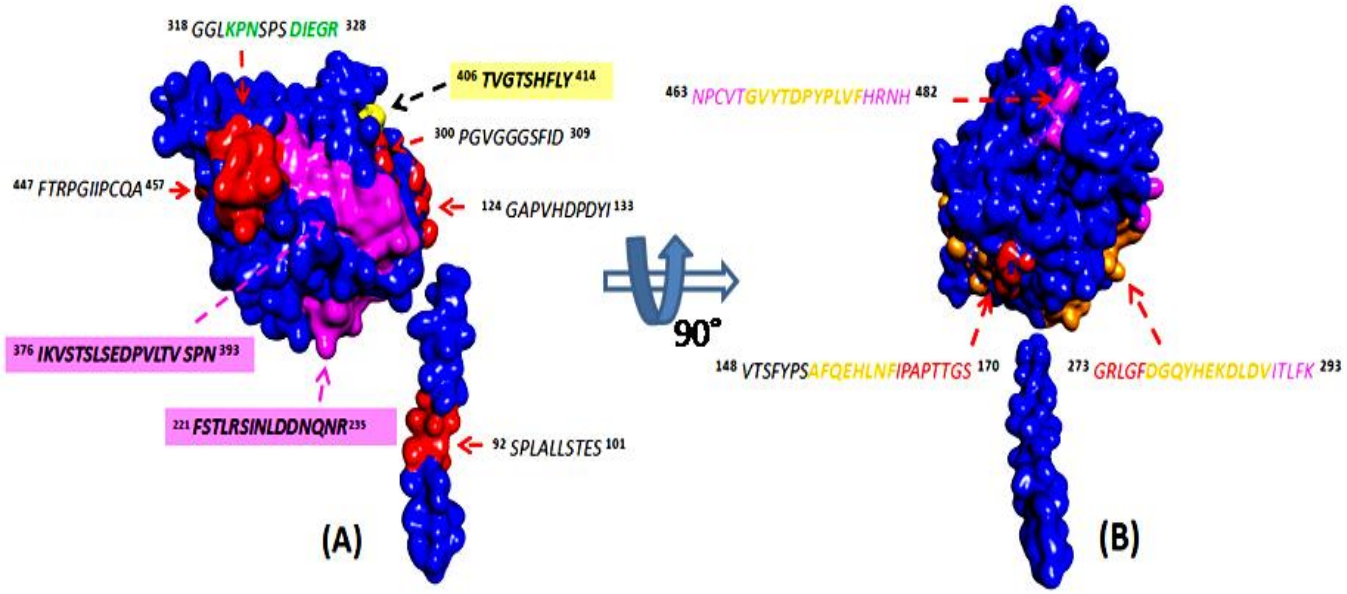


Fig. 4 : Position of predicted B and T cells epitopes in 3D structure of HN protein: (A) Non-overlapping epitopes: B cells epitopes = red, $TCD8^+$ cell epitopes= Yellow and $TCD4^+$ cell epitopes = purple. (B): Overlapping epitopes: the fragments of overlapping epitopes are colored in beige whereas the other non-overlapping fragments are colored according to their colors in (A).

ISSN No:-2456-2165

IV. DISCUSSION

Vaccination is one of the most widely used methods in the prevention of infectious diseases [65]. In the case of ND, conventional vaccines are available on the market and are commonly used in field [66]. These vaccines are usually formulated with attenuated or inactivated I-2, La Sota and Mukteswar strains [10]. These vaccine strains belong to the genotypes I, II and III, respectively [10,67]. Despite the continued utilization of these vaccines, ND outbreaks causing poultry morbidity and mortality are still observed in many countries, including Madagascar [25,68,69]. According to Miller and colleagues, commercially available ND vaccines can offer complete protection against clinical signs of the disease, but they are not able to completely prevent infection and spread of the virus. This decreased efficacy is due to the genetic distance between the circulating and vaccine strains [70,71]. Thus, the use of vaccines formulated with an antigen similar to the circulating strains has been shown to be effective in eliminating these obstacles [72,73].

In this study, we used bioinformatics methods (including sequences alignment, molecular modeling, immuno-informatics study and molecular docking) to determine the antigenic epitopes present in the HN protein of 12 NDV strains isolated in Madagascar between 1992 and 2011 [18,40]. The HN protein is chosen as a target because it is the major antigen carrier of the virus and could induce neutralizing antibodies during infection [38]. Vaccine immunity involves two types of immune responses, namely cell-mediated immunity dependent on $TCD4^+$ and $TCD8^+$ lymphocytes and humoral adaptive immunity involving B cells [74]. In effect, we predicted both groups of epitopes, including B and T-cell epitopes [26,75]. The results obtained in this study could be used in the design of a new multi-epitope vaccine to improve the control of ND.

In the design of a multi-epitope vaccine, it is preferable to use conserved epitopes, in order to obtain broad-spectrum protection and to have cross-immunity [76,77]. Accordingly, we first detected by multiple sequence alignment all conserved regions in the 12 sequences of HN. This bioinformatics analysis also provided insight into the positions of mutations in the HN protein of NDV isolates from Madagascar. As a result, we found 40 mutations and 14 conserved regions (Fig. 1). The impact of these mutations on the functions of HN protein is not yet well documented, with the exception of the change of amino acid glutamate (E) to lysine (K) at position 347 in the linear epitope 341-355 [78,79]. This mutation could increase the neuraminidase activity of the protein and enhance the release of the virus from infected cells [79]. Thus, further study should be required to analyze the effects of the others mutations in the structure and function of the HN protein.

After the detection of conserved regions, we predicted the B and T cells epitopes in the 12 HN proteins analyzed. Several immuno-informatics tools were available for such prediction [80-82]. However, we chose the tools ABCpred [42], BEPIPRED [43], DISCOTOPE 2.0 [44] and IEDB [45]. These first three servers were used for the prediction of B-cell epitopes while the last one is required for the screening of T-cells epitopes [83]. As results, we predicted 191 B-cells epitopes (157 linear and 34 discontinuous) and 199 T-cells epitopes (109 $TCD8^+$ and 90 $TCD4^+$). From these 390 predicted epitopes, we selected peptides that fell within the 14 conserved regions of the HN protein (Fig. 1). As a result, we retained 155 conserved epitopes, including 69 linear B-cell epitopes, 14 discontinuous B-cell epitopes, 37 $TCD8^+$ cell epitopes, and 35 $TCD4^+$ cell epitopes (Supplementary Tables 1, 2, 3 and 4).

In addition to conservation, an ideal epitope for multi-epitope vaccine design must also be antigenic [84] and accessible to an antibody or cell surface receptor [85]. Guarding these ideas, we further evaluated the antigenicity and membrane topology of these 155 conserved epitopes, using two servers immuno-informatics such VaxiJen v2.0 and TMHMM v2.0. After analysis, we found 102 epitopes (46 linear B-cell epitopes, 6 discontinuous B-cell epitopes, 26 $TCD8^+$ cell epitopes, and 24 $TCD4^+$ cell epitopes) having a VaxiJen v2.0 score higher than 0.4 (Supplementary Tables 1, 2, 3 and 4). Among these, only 45 peptides (14 linear B-cell epitopes, 3 discontinuous B-cell epitopes, 14 $TCD8^+$ cell epitopes, and 14 $TCD4^+$ cell epitopes) were localized in the exo-membrane regions of the HN protein. Therefore, these 45 epitopes were all retained in the lists of predicted epitopes in this study (Table 2, 3, 4, and 5).

It is well known that T-cell immune responses are more durable than B-cell responses, where antibody memory can be easily lost [86]. Therefore, many researchers have preferred to choose T epitopes than B epitope in the design of multi-epitope vaccines [87-89]. Unfortunately, the prediction of T-cell epitopes by immuno-computational tools has encountered many obstacles, which have plagued the success of vaccine design. The most important problem is the lack of chicken MHC alleles in the bioinformatics databases [86,90,91]. However, some studies have reported that chicken MHC B-F and B-L alleles have high similarity to human MHC or HLA [92,93]. For this reason, researchers have attempted to use HLA alleles (*HLA-A, HLA-B, HLA-C, HLA-DRB, and HLA-DQ*) in the prediction of T cell epitopes [26,94,95]. In the present study, we also used the HLA alleles to predict the epitope present in the HN of NDV isolated in Madagascar. As a result, we found that the 14 $TCD8^+$ cell epitopes and 13 $TCD4^+$ cell epitopes selected have interacted with *HLA-A*02:01/HLA-B*35:01* and *HLA-DRB1*03:01/HLA-DRB3*01:01* alleles respectively (Table 4 and 5). In contrast, Osman and his colleagues founded that the conserved $TCD8^+$ cell epitopes of NDV HN protein interacted with the alleles *HLA-B*44:02, HLA-B*15: 01, HLA-A*68:02, HLA-B*44:02, HLA-B*46:01, HLA-B*15:01* and $TCD4^+$ cell epitopes with *HLA-DRB1, HLA-DRB3, HLA-DRB5, HLA-DPA1/HLA-DPB1 and DQA1/DQB1* [91]. These contradictory results could be explained by the presence of mutations in the NDV HN protein belonging to genotype XI [12], which could the ability of certain peptide fragment to interact with the HLA alleles.

To ensure the accuracy of our predicting results, we docked the 28 predicted T-cells epitopes with two chicken MHC alleles namely *BF2*2101* and *BL2*019:01*.

ISSN No:-2456-2165

According to Koch and colleagues, the *BF2*2101* allele (*pdb ID = 3BEW*) had a large binding site suitable for many peptides and is present only in chickens [96]. In contrast, the *BL2*019:01* allele (*pdb ID= 6KVM*) is the only MHC class II allele from chicken that had a crystallized structure in the RCSB PDB database [97]. Therefore, we have selected them as receptors during molecular docking. The co-crystallized peptides are firstly removed from their receptors and then re-docked using the Autodock vina program [56] included in *AUTODOCK TOOLS* [57]. Complexes with lower interaction energy were considered as reference. The interaction energies of *BF2*2101* and *BL2*019:01* receptor with their co-crystallized ligands taken as references are -8.2 and -4.8 Kcal/mol respectively (Table 6 and 7). Comparing the docking results of our predicted epitopes with these values, we retained 21 epitopes including 9 $TCD8^+$ and 13 $TCD4^+$ cells epitopes (Table 6 and 7). In other words, five epitopes (*[86]KQ**VALESPLAL**LSTE[100]*, *[451]**TLFGEF**RIV[459]*, *[474]YPLVFHRNH[482]*, *[498]RLNPVSAVF[506]* *and* *[545]**IAEISNTLF**[554]*) were eliminated from the list of T-cells epitopes because their interaction energies were slightly higher compared to the reference values. In contrast, Osman and colleagues retained three of these eliminated epitopes (*[88]**VALESPLAL**[96]*, *[546]**AEISNTLF**G[554] et [548]**ISNTLF**GEF[556]*) after performing molecular docking with both *BF2*2101* and *BF2*04:01* alleles of chicken MHC class I [91]. We suggest that these discordances may be due to the types of MHC alleles and the difference in the size of the grid box used during docking. These factors may alter the interaction energy between ligands and receptors [98].

Overall, we identified 39 conserved epitopes (14 continuous B-cell epitopes, 3 discontinuous B-cell epitopes, 9 $TCD8^+$ cell epitopes, and 13 $TCD4^+$ cell epitopes) in the 12 proteins sequences analyzed (Table 2, 3, 6 and 7). However, some overlaps were still found between these retained epitopes (Table 8). In this regard, several researchers have previously reported that overlapping B and T cells epitopes are more advantageous in epitope-based vaccine design because they can generate not only a good antibody response, but also appropriate cellular responses and T cell memory [99-101]. But again, some said that the overlapping epitopes of $TCD8^+$ and $TCD4^+$ cells are likely to be candidates for multi-epitope vaccine design [102,103]. Thus, the observed overlaps in our predicted epitopes enhance the robustness of our results. Therefore, they are exploitable despite its primacy. Nevertheless, additional forward validation by other methods of epitope analysis is also needed.

To visualize these 39 epitopes predicted in the HN structure, we have firstly modeled the 3D structure of this protein in SWISS MODEL server [58]. We obtained a good model structure (Fig. 3A and B), identical to the crystal structure of HN from Australia Victoria strain (*pdb ID= 3t1e*) [59]. After that, we used this model for visualization of predicted epitope in D.S v2017 [53]. At first look, we found that all predicted epitopes (except for the linear B-cell epitope *[92]SPLALLSTES[101]*) are localized in the exo-membrane region of this protein, at its globular head (Fig. 4A and B). Jin and his collaborators have also reported that the immuno-dominant epitopes (IDEs) of HN are localized at the globular head of the protein [104]. But again, we also found that these 39 predicted epitopes are overlapped and formed 11 antigenic sites (Fig. 4 A and B).

In 1989, Iorio and colleagues employed monoclonal antibodies (mAbs) to predict the antigenic epitope of NDV HN and they found seven antigenic sites [105]. In 2015, Tao Li and his colleagues still performed *in vivo* study on several antigenic domains of NDV HN protein [106]. As a result, they reported that the peptide P2 (53-192aa) was one of the dominant linear antigenic domains of this protein [106]. Recently, Jin and his collaborators used the PepSCan method to determine the imuno-dominant epitopes (IDEs) of the NDV HN protein. For this, they identified five IDEs including: IDE1 (554-568aa), IDE2 (283-297aa), IDE3 (119-133aa) IDE4 (242-256aa) and IDE5 (328-342aa) [38]. In the current study, we used bioinformatics methods to predict the conserved epitopes of the HN protein of different NDV strains from Madagascar. A total of 39 epitopes were identified and they localized in 11 antigenic sites. Surprisingly, the five linear B-cell epitopes (*[92]SPLALLSTES[101] ; [124]GAPVHDPDYI[133] ; [155]AFQEHLNFIP[164]; [158]EHLNFIPAPT[167]* *and* *[161]NFIPAPTTGS[170]*) and the two $TCD8^+$ epitopes (*[148] VTSFYPSAF [156] and [154] SAFQEHLNF [162]*) predicted in this study are located in the region flanked by the peptide P2 designated by Tao Li and his collaborators [106]. In addition, the three linear B-cell epitopes (*[277]FDGQYHEKDL[286]*, *[278]DGQYHEKDL[287]* and *[279]GQYHEKDLDV[288]*) are overlapped in the region of IDE2 reported by Jin and colleagues [104]. Moreover, the two linear B-cell epitopes (*[318]GGLKPNSPSD[327]* and *[319]GLKPSSPSDI[328]*) and the three discontinuous B-cell epitopes (*[321]KPNSP*DIA*EG[332]; [322]PNSP**IA*EGR[333]* and *[322]PNSP*DIA*EG[332]*) predicted overlapped with IDE5 [104]. In one direction, these results suggested that the methods that we had employed in this research are suitable and can replace the empirical methods of epitopes analysis [33,34]. This means that bioinformatics and immuno-informatics techniques provide a promising and cost-effective approach to identify potential T and B-cell epitopes [80]. But in the other direction, these results also revealed the presence of epitopes and/or antigenic sites in the HN protein that were no more identified by other researchers. Thus, the methods and the results obtained in this study could be used as a reference in prediction of antigenic epitopes of other NDV proteins and viral strains.

V. CONCLUSION

Using bioinformatics methods, we identified 39 conserved epitopes (14 continuous B-cell epitopes, 3 discontinuous B-cells epitopes, 9 $TCD8^+$ cell epitopes, and 13 $TCD4^+$ cell epitopes) in the HN protein of 12 NDV isolates from Madagascar. These epitopes are generally localized at the globular head of the protein and overlap in 11 antigenic sites. Firstly, our results are comparable with the results obtained by other researchers who use *in vitro* and/or *in vivo* methods. Thus, they can be used for the development of a novel multi-epitopic vaccine against ND. Additionally, these results revealed the presence of other antigenic sites in the HN protein that have not been

identified by other researchers. In effect, this study could be used as a reference in future studies on epitope analysis in NDV. As this is a preliminary study, there is also necessary to perform further *in vitro* and/or *in vivo* analyses to improve the reliability of these results. There is also a need to explore the antigenic properties of other viral proteins in NDV to expand the spectrum of vaccine efficacy that we thought to design.

ACKNOWLEDGEMENTS

The authors thank the Malagasy Institute of Veterinary Vaccines (IMVAVET) for its support in research and technology.

REFERENCES

[1.] Dimitrov KM, Abolnik C, Afonso CL, Albina E, Bahl J, et al. (2019) Updated unified phylogenetic classification system and revised nomenclature for Newcastle disease virus. Infect Genet Evol 74: 103917.
[2.] Amarasinghe GK, Ayllón MA (2019) Taxonomy of the order Mononegavirales: update 2019. 164: 1967-1980.
[3.] Aldous EW (2001) Detection and differentiation of Newcastle disease virus (avian paramyxovirus type 1). Avian pathology v. 30: pp. 117-128-2001 v.2030 no.2002.
[4.] Alexander DJ (1995) The epidemiology and control of avian influenza and Newcastle disease. J Comp Pathol 112: 105-126.
[5.] Ganar K, Das M, Sinha S, Kumar S (2014) Newcastle disease virus: current status and our understanding. Virus Res 184: 71-81.
[6.] Annapragada A, Borgerson C, Iams S, Ravelomanantsoa MA, Crawford GC, et al. (2019) Modeling the Impact of Newcastle Disease Virus Vaccinations on Chicken Production Systems in Northeastern Madagascar. Frontiers in Veterinary Science 6.
[7.] Amoia CFANG, Nnadi PA, Ezema C, Couacy-Hymann E (2021) Epidemiology of Newcastle disease in Africa with emphasis on Côte d'Ivoire: A review. Veterinary world 14: 1727-1740.
[8.] Olivier Fridolin M (2013) Épidémiologie de la maladie de Newcastle en aviculture villageoise à Madagascar.
[9.] Randriamamisolonirina NT (2017) CMH des Poulets et Resistance aux Maladies Virales à Madagascar : Cas de la Maladie de Newcastle. Faculté des Sciences: Université d'Antananarivo. 85 p.
[10.] Dimitrov KM, Ramey AM, Qiu X, Bahl J, Afonso CL (2016) Temporal, geographic, and host distribution of avian paramyxovirus 1 (Newcastle disease virus). Infection, genetics and evolution 39: 22-34.
[11.] Bello MB, Yusoff K, Ideris A, Hair-Bejo M, Peeters BPH, et al. (2018) Diagnostic and Vaccination Approaches for Newcastle Disease Virus in Poultry: The Current and Emerging Perspectives. BioMed Research International 2018: 7278459.
[12.] Maminiaina OF (2011) Caractérisation des virus de la maladie de Newcastle (APMV-1), circulant sur les hautes terres de Madagascar. 209-209 p.
[13.] Mohamed MH, Abdelaziz AM, Kumar S, Al-Habib MA, Megahed MM (2016) Effect of phylogenetic diversity of velogenic Newcastle disease virus challenge on virus shedding post homologous and heterologous DNA vaccination in chickens. Avian Pathol 45: 228-234.
[14.] Wajid A, Basharat A, Bibi T, Rehmani SF (2018) Comparison of protection and viral shedding following vaccination with Newcastle disease virus strains of different genotypes used in vaccine formulation. Trop Anim Health Prod 50: 1645-1651.
[15.] Rauw F, Gardin Y, van den Berg T, Lambrecht B (2009) La vaccination contre la maladie de Newcastle chez le poulet (Gallus gallus). Biotechnology, Agronomy and Society and Environment 13: 587-596.
[16.] Miller PJ, King DJ, Afonso CL, Suarez DL (2007) Antigenic differences among Newcastle disease virus strains of different genotypes used in vaccine formulation affect viral shedding after a virulent challenge. Vaccine 25: 7238-7246.
[17.] Miller PJ, Estevez C, Yu Q, Suarez DL, King DJ (2009) Comparison of viral shedding following vaccination with inactivated and live Newcastle disease vaccines formulated with wild-type and recombinant viruses. Avian diseases 53: 39-49.
[18.] Maminiaina OF, Gil P, Albina E, Lancelot R, Andriamanivo HR, et al. (2010) Newcastle Disease Virus in Madagascar : Identification of an Original Genotype Possibly Deriving from a Died Out Ancestor of Genotype IV. PLoS ONE 5: 12-12.
[19.] S. RM, F. MO, Andriantsimahavandy Abel A., GIL Patricia, Almeida RSd, et al. (2020) Isolement et identification du virus de la maladie de Newcastle à Madagascar. In: Scientifique MdlESedlR, editor: Centre d'Information et de Documentation Scientifique et Technique. pp. 53-60.
[20.] Annapragada A, Borgerson C, Iams S, Ravelomanantsoa MA, Crawford GC, et al. (2019) Modeling the Impact of Newcastle Disease Virus Vaccinations on Chicken Production Systems in Northeastern Madagascar. Frontiers in veterinary science 6: 305-305.
[21.] Sabra M, Dimitrov KM, Goraichuk IV, Wajid A, Sharma P, et al. (2017) Phylogenetic assessment reveals continuous evolution and circulation of pigeon-derived virulent avian avulaviruses 1 in Eastern Europe, Asia, and Africa. BMC Veterinary Research 13: 291.
[22.] da Silva AP, Aston EJ, Chiwanga GH, Birakos A, Muhairwa AP, et al. (2020) Molecular Characterization of Newcastle Disease Viruses Isolated from Chickens in Tanzania and Ghana. Viruses 12: 916.
[23.] Enayatkhani M, Hasaniazad M, Faezi S (2021) Reverse vaccinology approach to design a novel multi-epitope vaccine candidate against COVID-19: an in silico study. 39: 2857-2872.

[24.] Jahangirian E, Jamal GA, Nouroozi M, Mohammadpour A (2021) A reverse vaccinology and immunoinformatics approach for designing a multiepitope vaccine against SARS-CoV-2. Immunogenetics 73: 459-477.
[25.] Omony JB, Wanyana A, Mugimba KK, Kirunda H, Nakavuma JL, et al. (2021) Epitope Peptide-Based Predication and Other Functional Regions of Antigenic F and HN Proteins of Waterfowl and Poultry Avian Avulavirus Serotype-1 Isolates From Uganda. Frontiers in Veterinary Science 8.
[26.] Raza A, Asif Rasheed M, Raza S, Tariq Navid M, Afzal A, et al. (2022) Prediction and analysis of multi epitope based vaccine against Newcastle disease virus based on haemagglutinin neuraminidase protein. Saudi Journal of Biological Sciences.
[27.] Behbahani M, Moradi M, Mohabatkar H (2021) In silico design of a multi-epitope peptide construct as a potential vaccine candidate for Influenza A based on neuraminidase protein. In silico pharmacology 9: 36-36.
[28.] Yusoff K, Nesbit M, McCartney H, Meulemans G, Alexander DJ, et al. (1989) Location of neutralizing epitopes on the fusion protein of Newcastle disease virus strain Beaudette C. J Gen Virol 70 (Pt 11): 3105-3109.
[29.] Yusoff K, Nesbit M, McCartney H, Emmerson PT, Samson AC (1988) Mapping of three antigenic sites on the haemagglutinin-neuraminidase protein of Newcastle disease virus. Virus Res 11: 319-333.
[30.] Panshin A, Shihmanter E, Weisman Y, Orvell C, Lipkind M (1997) Antigenic epitope characterization of matrix protein of Newcastle disease virus using monoclonal antibody approach: contrasting variability amongst NDV strains. Comp Immunol Microbiol Infect Dis 20: 177-189.
[31.] Mebatsion T, Koolen MJM, de Vaan LTC, de Haas N, Braber M, et al. (2002) Newcastle disease virus (NDV) marker vaccine: an immunodominant epitope on the nucleoprotein gene of NDV can be deleted or replaced by a foreign epitope. Journal of virology 76: 10138-10146.
[32.] Borley DW, Mahapatra M, Paton DJ, Esnouf RM, Stuart DI, et al. (2013) Evaluation and Use of In-Silico Structure-Based Epitope Prediction with Foot-and-Mouth Disease Virus. PLOS ONE 8: e61122.
[33.] Moutaftsi M, Peters B, Pasquetto V, Tscharke DC, Sidney J, et al. (2006) A consensus epitope prediction approach identifies the breadth of murine TCD8+-cell responses to vaccinia virus. Nature Biotechnology 24: 817-819.
[34.] Guedes RLM, Rodrigues CMF, Coatnoan N, Cosson A, Cadioli FA, et al. (2019) A comparative in silico linear B-cell epitope prediction and characterization for South American and African Trypanosoma vivax strains. Genomics 111: 407-417.
[35.] Rawal K, Sinha R, Abbasi BA, Chaudhary A, Nath SK, et al. (2021) Identification of vaccine targets in pathogens and design of a vaccine using computational approaches. Scientific Reports 11: 17626.
[36.] Majid M, Andleeb S (2019) Designing a multi-epitopic vaccine against the enterotoxigenic Bacteroides fragilis based on immunoinformatics approach. Scientific Reports 9: 19780.
[37.] Lohia N, Baranwal M (2020) An Immunoinformatics Approach in Design of Synthetic Peptide Vaccine Against Influenza Virus. Methods Mol Biol 2131: 229-243.
[38.] Jin Z, Wei Q, Bi Y, Li Y, Huo N, et al. (2021) Identification of a potential neutralizing linear epitope of hemagglutinin-neuraminidase in Newcastle disease virus. Virology Journal 18: 8.
[39.] Guex N, Peitsch MC (1997) SWISS-MODEL and the Swiss-Pdb Viewer: An environment for comparative protein modeling. ELECTROPHORESIS 18: 2714-2723.
[40.] Almeida RSd, Maminiaina OF, Gil P, Hammoumi S, Molia S, et al. (2009) Africa, a reservoir of new virulent strains of Newcastle disease virus? Amsterdam: Elsevier. pp. 3127-3129.
[41.] Sanchez-Trincado JL, Gomez-Perosanz M, Reche PA (2017) Fundamentals and methods for T-and B-cell epitope prediction. Journal of immunology research 2017.
[42.] Saha S, Raghava GPS (2006) Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. Proteins: Structure, Function, and Bioinformatics 65: 40-48.
[43.] Jespersen MC, Peters B, Nielsen M, Marcatili P (2017) BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes. Nucleic Acids Research 45: W24-W29.
[44.] Kringelum JV, Lundegaard C, Lund O, Nielsen M (2012) Reliable B Cell Epitope Predictions: Impacts of Method Development and Improved Benchmarking. PLOS Computational Biology 8: e1002829-e1002829.
[45.] Kim Y, Ponomarenko J, Zhu Z, Tamang D, Wang P, et al. (2012) Immune epitope database analysis resource. Nucleic Acids Research 40: W525-W530.
[46.] Reynisson B, Alvarez B, Paul S, Peters B, Nielsen M (2020) NetMHCpan-4.1 and NetMHCIIpan-4.0: improved predictions of MHC antigen presentation by concurrent motif deconvolution and integration of MS MHC eluted ligand data. Nucleic Acids Res 48: W449-w454.
[47.] Trolle T, McMurtrey CP, Sidney J, Bardet W, Osborn SC, et al. (2016) The Length Distribution of Class I–Restricted T Cell Epitopes Is Determined by Both Peptide Supply and MHC Allele–Specific Binding Preference. The Journal of Immunology 196: 1480.
[48.] Weiskopf D, Angelo MA, de Azeredo EL, Sidney J, Greenbaum JA, et al. (2013) Comprehensive analysis of dengue virus-specific responses supports an HLA-linked protective role for CD8+ T cells. Proc Natl Acad Sci U S A 110: E2046-2053.
[49.] Wang P, Sidney J, Kim Y, Sette A, Lund O, et al. (2010) Peptide binding predictions for HLA DR, DP and DQ molecules. BMC Bioinformatics 11: 568.
[50.] Fleri W, Paul S, Dhanda SK, Mahajan S, Xu X, et al. (2017) The Immune Epitope Database and Analysis

Resource in Epitope Discovery and Synthetic Vaccine Design. Front Immunol 8: 278.
[51.] Doytchinova IA, Flower DR (2007) VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. BMC Bioinformatics 8: 4-4.
[52.] Krogh A, Larsson B, von Heijne G, Sonnhammer EL (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol 305: 567-580.
[53.] Biovia DS (2017) Discovery studio modeling environment. Release.
[54.] Lamiable A, Thévenet P, Rey J, Vavrusa M, Derreumaux P, et al. (2016) PEP-FOLD3: faster de novo structure prediction for linear peptides in solution and in complex. Nucleic Acids Res 44: W449-454.
[55.] Maupetit J, Derreumaux P, Tufféry P (2010) A fast method for large-scale de novo peptide and miniprotein structure prediction. J Comput Chem 31: 726-738.
[56.] Trott O, Olson AJ (2010) AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. Journal of Computational Chemistry 31: 455-461.
[57.] Goodsell DS, Morris GM, Olson AJ (1996) Automated docking of flexible ligands: Applications of autodock. Journal of Molecular Recognition 9: 1-5.
[58.] Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, et al. (2018) SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res 46: W296-w303.
[59.] Welch BD, Yuan P, Bose S, Kors CA, Lamb RA, et al. (2013) Structure of the Parainfluenza Virus 5 (PIV5) Hemagglutinin-Neuraminidase (HN) Ectodomain. PLOS Pathogens 9: e1003534.
[60.] Ilyas U, Ilyas U, Zaman Su, Syed A, Afzaal H, et al. (2018) ADME/T Prediction, Molecular Docking, and Biological Screening of 1,2,4-Triazoles as Potential Antifungal Agents. Journal of Applied Bioinformatics & Computational Biology 07.
[61.] Hooft RWW, Sander C, Vriend G (1997) Objectively judging the quality of a protein structure from a Ramachandran plot. Bioinformatics 13: 425-430.
[62.] Tan WS, Lau CH, Ng BK, Ibrahim AL, Yusoff K (1995) Nucleotide sequence of the haemagglutinin-neuraminidase (HN) gene of a Malaysian heat resistant viscerotropic-velogenic Newcastle disease virus (NDV) strain AF2240. DNA Seq 6: 47-50.
[63.] Yusoff K, Tan WS, Lau CH, Ng BK, Ibrahim AL (1996) Sequence of the haemagglutinin-neuraminidase gene of the Newcastle disease virus oral vaccine strain V4(UPM). Avian Pathol 25: 837-844.
[64.] Yusoff K, Tan WS (2001) Newcastle disease virus: Macromolecules and opportunities. Avian pathology : journal of the WVPA 30: 439-455.
[65.] Bello MB, Yusoff K, Ideris A, Hair-Bejo M, Jibril AH, et al. (2020) Exploring the Prospects of Engineered Newcastle Disease Virus in Modern Vaccinology. Viruses 12: 451.
[66.] Alders R, Spradbrow P (2000) La maladie de Newcastle dans les élevages avicoles villageois. International Network for Family Poultry Development 70.
[67.] Izquierdo-Lara R, Chumbe A, Calderón K, Fernández-Díaz M, Vakharia VN (2019) Genotype-matched Newcastle disease virus vaccine confers improved protection against genotype XII challenge: The importance of cytoplasmic tails in viral replication and vaccine design. PloS one 14: e0209539-e0209539.
[68.] Abd El-Hamid HS, Shafi ME, Albaqami NM, Ellakany HF, Abdelaziz NM, et al. (2020) Sequence analysis and pathogenicity of Avian Orthoavulavirus 1 strains isolated from poultry flocks during 2015-2019. 16: 253.
[69.] Ogali IN, Okumu PO, Mungube EO, Lichoti JK, Ogada S, et al. (2020) Genomic and Pathogenic Characteristics of Virulent Newcastle Disease Virus Isolated from Chicken in Live Bird Markets and Backyard Flocks in Kenya. International Journal of Microbiology 2020: 4705768.
[70.] Miller PJ, Afonso CL, El Attrache J, Dorsey KM, Courtney SC, et al. (2013) Effects of Newcastle disease virus vaccine antibodies on the shedding and transmission of challenge viruses. Dev Comp Immunol 41: 505-513.
[71.] Miller PJ, King DJ, Afonso CL, Suarez DL (2007) Antigenic differences among Newcastle disease virus strains of different genotypes used in vaccine formulation affect viral shedding after a virulent challenge. Vaccine 25: 7238-7246.
[72.] Bello MB, Mahamud SN, Yusoff K, Ideris A, Hair-Bejo M, et al. (2020) Development of an Effective and Stable Genotype-Matched Live Attenuated Newcastle Disease Virus Vaccine Based on a Novel Naturally Recombinant Malaysian Isolate Using Reverse Genetics. Vaccines 8.
[73.] Sultan HA, Elfeil WK, Nour AA, Tantawy L, Kamel EG, et al. (2021) Efficacy of the Newcastle Disease Virus Genotype VII.1.1-Matched Vaccines in Commercial Broilers. Vaccines 10: 29.
[74.] Kapczynski DR, Afonso CL, Miller PJ (2013) Immune responses of poultry to Newcastle disease virus. Developmental & Comparative Immunology 41: 447-453.
[75.] Saadi M, Karkhah A, Nouri HR (2017) Development of a multi-epitope peptide vaccine inducing robust T cell responses against brucellosis using immunoinformatics based approaches. Infect Genet Evol 51: 227-234.
[76.] Eickhoff CS, Terry FE, Peng L, Meza KA, Sakala IG, et al. (2019) Highly conserved influenza T cell epitopes induce broadly protective immunity. Vaccine 37: 5371-5381.
[77.] Staneková Z, Varečková E (2010) Conserved epitopes of influenza A virus inducing protective immunity and their prospects for universal vaccine development. Virol J 7: 351.
[78.] Gotoh B, Sakaguchi T, Nishikawa K, Inocencio NM, Hamaguchi M, et al. (1988) Structural features unique to each of the three antigenic sites on the

hemagglutinin-neuraminidase protein of Newcastle disease virus. Virology 163: 174-182.
[79.] Cho SH, Kwon HJ, Kim TE, Kim JH, Yoo HS, et al. (2008) Characterization of a recombinant Newcastle disease virus vaccine strain. Clin Vaccine Immunol 15: 1572-1579.
[80.] El-Manzalawy Y, Dobbs D, Honavar VG (2017) In Silico Prediction of Linear B-Cell Epitopes on Proteins. Methods in molecular biology (Clifton, NJ) 1484: 255-264.
[81.] Sun P, Guo S, Sun J, Tan L, Lu C, et al. (2019) Advances in In-silico B-cell Epitope Prediction. Curr Top Med Chem 19: 105-115.
[82.] Chen H-Z, Tang L-L, Yu X-L, Zhou J, Chang Y-F, et al. (2020) Bioinformatics analysis of epitope-based vaccine design against the novel SARS-CoV-2. Infectious Diseases of Poverty 9: 88.
[83.] Sanchez-Trincado JL, Gomez-Perosanz M, Reche PA (2017) Fundamentals and Methods for T- and B-Cell Epitope Prediction. Journal of Immunology Research 2017: 2680160.
[84.] Bibi S, Ullah I, Zhu B, Adnan M, Liaqat R, et al. (2021) In silico analysis of epitope-based vaccine candidate against tuberculosis using reverse vaccinology. Scientific Reports 11: 1249.
[85.] Caoili SEC (2010) B-cell epitope prediction for peptide-based vaccine design: towards a paradigm of biological outcomes. Proceedings of the First ACM International Conference on Bioinformatics and Computational Biology. Niagara Falls, New York: Association for Computing Machinery. pp. 602–610.
[86.] Mustafa Badawi M, AbdElhamid Fadl Alla A, Sleak Alam S, Aljack Mohamed W, Adil Nasr-Eldin Osman D, et al. (2022) Immunoinformatics Predication and in silico Modeling of Epitope-Based Peptide Vaccine Against virulent Newcastle Disease Viruses. American Journal of Infectious Diseases and Microbiology 4: 61-71.
[87.] Oany AR, Emran A-A, Jyoti TP (2014) Design of an epitope-based peptide vaccine against spike protein of human coronavirus: an in silico approach. Drug design, development and therapy 8: 1139-1149.
[88.] Adam KM (2021) Immunoinformatics approach for multi-epitope vaccine design against structural proteins and ORF1a polyprotein of severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2). Tropical Diseases, Travel Medicine and Vaccines 7: 22.
[89.] Mugunthan SP, Harish MC (2021) Multi-epitope-Based Vaccine Designed by Targeting Cytoadherence Proteins of Mycoplasma gallisepticum. ACS omega 6: 13742-13755.
[90.] Ali SA, Almofti YA, Abd-elrahman KA (2019) Immunoinformatics Approach for Multiepitopes Vaccine Prediction against Glycoprotein B of Avian Infectious Laryngotracheitis Virus. Advances in Bioinformatics 2019: 1270485.
[91.] Mohamed Osman M, Elamin E, Al-Nour M, Alam S, Adam R, et al. (2016) In Silico Design of Epitope Based Peptide Vaccine against Virulent Strains of HN- Newcastle Disease Virus (NDV) in Poultry Species.
[92.] Vainio O, Koch C, Toivanen A (1984) B-L antigens (class II) of the chicken major histocompatibility complex control T-B cell interaction. Immunogenetics 19: 131-140.
[93.] Hála K, Boyd R, Wick G (1981) Chicken major histocompatibility complex and disease. Scand J Immunol 14: 607-616.
[94.] Hou Y, Guo Y, Wu C, Shen N, Jiang Y, et al. (2012) Prediction and identification of T cell epitopes in the H5N1 influenza virus nucleoprotein in chicken. PLoS One 7: e39344.
[95.] Badawi MM, Fadl Alla A, Alam SS, Mohamed WA, Osman D, et al. (2016) Immunoinformatics predication and in silico modeling of epitope-based peptide vaccine against virulent Newcastle disease viruses. Am J Infect Dis Microbiol 4: 61-71.
[96.] Koch M, Camp S, Collen T, Avila D, Salomonsen J, et al. (2007) Structures of an MHC class I molecule from B21 chickens illustrate promiscuous peptide binding. Immunity 27: 885-899.
[97.] Zhang L, Li X, Ma L, Zhang B, Meng G (2020) A Newly Recognized Pairing Mechanism of the α- and β-Chains of the Chicken Peptide-MHC Class II Complex. 204: 1630-1640.
[98.] Ivonie U, Yanuar A, Firdayani (2018) VIRTUAL SCREENING OF INDONESIAN HERBAL DATABASE FOR CP ALLOSTERIC MODULATOR OF HEPATITIS B VIRUS. International Journal of Applied Pharmaceutics 10: 190-195.
[99.] Hosseini SS, Aghaiypour Kolyani K, Rafiei Tabatabaei R, Goudarzi H, Akhavan Sepahi A, et al. (2021) In silico prediction of B and T cell epitopes based on NDV fusion protein for vaccine development against Newcastle disease virus. Veterinary research forum : an international quarterly journal 12: 157-165.
[100.] Parida R, Shaila MS, Mukherjee S, Chandra NR, Nayak R (2007) Computational analysis of proteome of H5N1 avian influenza virus to define T cell epitopes with vaccine potential. Vaccine 25: 7530-7539.
[101.] Lal G, Shaila MS, Nayak R (2006) Activated mouse T cells downregulate, process and present their surface TCR to cognate anti-idiotypic CD4+ T cells. Immunol Cell Biol 84: 145-153.
[102.] Lohia N, Baranwal M (2014) Conserved peptides containing overlapping CD4+ and CD8+ T-cell epitopes in the H1N1 influenza virus: an immunoinformatics approach. Viral Immunol 27: 225-234.
[103.] Lohia N, Baranwal M (2015) Identification of Conserved Peptides Comprising Multiple T Cell Epitopes of Matrix 1 Protein in H1N1 Influenza Virus. Viral Immunol 28: 570-579.
[104.] Jin Z, Wei Q, Bi Y, Li Y, Huo N, et al. (2021) Identification of a potential neutralizing linear epitope of hemagglutinin - neuraminidase in Newcastle disease virus. Virology Journal: 1-11.
[105.] Iorio RM, Glickman RL, Riel AM, Sheehan JP, Bratt MA (1989) Functional and neutralization profile of seven overlapping antigenic sites on the HN

glycoprotein of Newcastle disease virus: monoclonal antibodies to some sites prevent viral attachment. Virus Res 13: 245-261.

[106.] Li T, Wang G, Shi B, Liu P, Si W, et al. (2015) Comprehensive Analysis and Characterization of Linear Antigenic Domains on HN Protein from Genotype VII Newcastle Disease Virus Using Yeast Surface Display System. PLOS ONE 10: e0131723.

ISSN No:-2456-2165

SUPPLEMENTARY DATA

N°	Start	End	Sequence	Length	Score VAXIJEN	TMHMM 2.0 score
1	92	101	SPLALLSTES	10	0.8382	outside
2	124	133	GAPVHDPDYI	10	0.4296	outside
3	155	164	AFQEHLNFIP	10	1.0412	outside
4	158	167	EHLNFIPAPT	10	1.3725	outside
5	161	170	NFIPAPTTGS	10	0.7259	outside
6	273	282	GRLGFDGQYH	10	1.6473	outside
7	274	283	RLGFDGQYHE	10	1.1843	outside
8	277	286	FDGQYHEKDL	10	1.1298	outside
9	278	287	DGQYHEKDLD	10	1.4379	outside
10	279	288	GQYHEKDLDV	10	1.5970	outside
11	300	309	PGVGGGSFID	10	0.7099	outside
12	318	327	GGLKPNSPSD	10	0.4094	outside
13	380	385	TSLGED	6	0.6376	outside
14	447	457	FTRPGIIPCQA	10	0.5708	outside
15	7	17	RVALENEEREA	11	0.8950	Inside
16	7	16	RVALENEERE	10	0.9370	Inside
17	8	16	VALENEERE	10	1.4286	Inside
18	8	18	VALENEEREAK	12	1.1673	Inside
19	9	17	ALENEEREA	9	1.0640	Inside
20	63	79	ITRAEERITSALSSNRD	17	0.5709	Inside
21	64	76	SRAEERITSTLRS	13	0.8261	Inside
22	71	80	TSALSSNRDV	10	0.4127	Inside
23	106	115	AITSLSYQIN	10	1.1126	Inside
24	107	116	ITSLSYQING	10	0.7171	Inside
25	165	174	APTTGSGCTR	10	0.4141	Inside
26	174	183	RIPSFDISAT	10	1.0609	Inside
27	181	190	SATHYCYTHN	10	0.7438	Inside
28	183	192	THYCYTHNVI	10	0.5864	Inside
29	196	201	CRDHSH	6	0.6068	Inside
30	209	218	GVLRTSATGR	10	0.4677	Inside
31	227	236	INLDDNQNRK	10	0.6303	Inside
32	227	235	INLDDNQNR	9	0.4529	Inside
33	232	241	NQNRKSCSVS	10	1.1210	Inside
34	235	244	RKSCSVSATP	10	1.2922	Inside
35	239	248	SVSATPLGCD	10	0.8021	Inside
36	240	249	VSATPLGCDI	10	0.7591	Inside
37	256	267	ETEEEDYRSVTP	12	0.9201	Inside
38	259	268	EEDYRSVTPT	10	0.9948	Inside
39	269	278	SMVHGRLGFD	10	1.7114	Inside
40	280	292	QYHEKDLDVITLF	13	0.7614	Inside
41	324	333	SPSDIAQEGR	10	0.9343	Inside
42	328	337	IAQEGRYVIY	10	0.7550	Inside
43	334	343	YVIYKRYNDT	10	0.6038	Inside
44	352	361	IRMAKSSYKP	10	0.9521	Inside
45	478	483	FHRNHT	6	0.5437	Inside

ISSN No:-2456-2165

46	516	523	RVSSGSTK	8	0.8544	Inside
47	82	91	DRVYKQVALE	10	0.1305	
48	87	96	QVALESPLAL	10	0.2804	
49	88	97	VALESPLALL	10	0.1601	
50	96	105	LLSTESVIMN	10	0.3069	
51	110	158	LSYQINGAANNSGCGAPVHDPDYIGGIGKELIVDDTSEVTSFYPSAFQE	49	0.3210	
52	111	159	SYQINGAANNSGCGAPVHDPDYIGGIGKELIVDDTSEVTSFYPSAFQEH	49	0.2745	
53	132	141	YIGGIGKELI	10	0.3631	
54	143	152	DDTSEVTSFY	10	0.2272	
55	150	159	SFYPSAFQEH	10	0.3535	
56	162	170	FIPAPTTGS	9	0.3499	
57	162	171	FIPAPTTGSG	10	0.1091	
58	163	170	IPAPTTGS	7	0.0235	
59	186	195	CYTHNVILSG	10	0.3173	
60	190	199	NVILSGCRDH	10	-0.2320	
61	194	203	SGCRDHSHSH	10	0.2244	
62	228	235	NLDDNQNR	8	0.1928	
63	280	300	QYHEKDLDVITLFKDWVANYP	21	0.1102	
64	283	292	EKDLDVITLF	10	0.3081	
65	288	297	VITLFKDWVA	10	-0.6661	
66	293	302	KDWVANYPGV	10	-0.4666	
67	321	330	KPNSPSDIAQ	10	0.2920	
68	380	385	TSLSED	6	0.2950	
69	433	457	NKTATLHSPYAFNAFTRPGTVPCQA	25	0.3543	

Supplementary Table 9 : Linear B-cell epitopes predict in this study

N°	Start	End	Sequence	Length	VaxiJen score	TMHMM 2.0 score
1	321	332	KPNSP*DIA*EG	12	0.9504	outside
2	322	333	PNSP**IA*EGR	12	0.6735	outside
3	322	332	PNSP*DIA*EG	11	0.8039	outside
4	255	266	TET*EE*YR*VT	12	0.5083	Inside
5	321	333	KPNSP*DIA*EGR	13	1.1298	Inside
6	322	333	PN****I**EGR	12	2.6624	Inside
7	111	115	SYQIN	5		
8	145	145	T	1		
9	198	198	D	1		
10	200	201	SH	2		
11	231	232	DN	2		
12	232	232	N	1		
13	255	265	T***EE**R*V	11		
14	259	265	EE**R*V	7		

Supplementary Table 10 : Discontinious B-cell epitopes predict in this study

ISSN No:-2456-2165

N°	Start	End	Peptide	Length	HLA Alleles	Percentile Rank	VaxiJen 2.0 score	TMHMM 2.0 Result
1	148	156	VTSFYPSAF	9	HLA-B*35:01 HLA-B*57:01	0.95 0.72	0.4192	outside
2	154	162	SAFQEHLNF	9	HLA-B*35:01 HLA-B*57:01	0.08 0.27	0.6244	outside
3	177	185	SFDISATHY	9	HLA-B*35:01	0.57	1.7398	outside
4	273	281	GRLGFDGQY	9	HLA-B*27:05	0.04	1.9611	outside
5	299	307	YPGVGGGSF	9	HLA-B*07:02 HLA-B*35:01	0.1 0.04	1.0615	outside
6	406	414	TVGTSHFLY	9	HLA-B*35:01	0.58	0.5265	outside
7	419	427	SYFSPALLY	9	HLA-B*35:01 HLA-B*27:05	0.56	0.6322	outside
8	468	476	GVYTDPYPL	9	HLA-A*02:01	0.54	0.6599	outside
9	470	478	YTDPYPLVF	9	HLA-A*02:01 HLA-B*07:02 HLA-B*35:01 HLA-B*57:01	0.56 0.92 0.08 0.2	1.0589	outside
10	474	482	YPLVFHRNH	9	HLA-B*35:01	0.77	1.1189	outside
11	498	506	RLNPVSAVF	9	HLA-A*02:01 HLA-B*07:02 HLA-B*57:01 HLA-B*35:01	0.91 0.76 0.53 0.56	1.4846	outside
12	545	553	IAEISNTLF	9	HLA-B*35:01	0.53	0.4915	outside
13	551	559	TLFGEFRIV	9	HLA-A*02:01	0.19	0.4567	outside
14	557	565	RIVPLLVEI	9	HLA-A*02:01	0.16	1.1852	outside
15	104	112	MNAITSLSY	9	HLA-B*35:01	0.33	0.7875	Inside
16	106	114	AITSLSYQI	9	HLA-A*02:01	1.0	1.1163	Inside
17	179	187	DISATHYCY	9	HLA-B*35:01	0.29	1.7254	Inside
18	262	270	YRSVTPTSM	9	HLA-B*27:05	0.11	0.9582	Inside
19	352	360	IRMAKSSYK	9	HLA-B*27:05	0.47	0.7352	Inside
20	362	370	GRFGGKRVQ	9	HLA-B*27:05	0.7	0.6496	Inside
21	396	404	TLMGAEGRV	9	HLA-A*02:01	0.86	0.7667	Inside
22	427	435	YPMIVNNKT	9	HLA-B*35:01	0.73	0.9121	Inside
23	476	484	LVFHRNHTL	9	HLA-A*02:01 HLA-B*07:02	0.65 0.72	0.8242	Inside
24	524	532	AAYTTSTCF	9	HLA-B*35:01	0.2	0.8479	Inside
25	533	541	KVVKTNRTY	9	HLA-B*35:01 HLA-B*57:01	0.83 0.38	0.5512	Inside
26	544	552	SIAEISNTL	9	HLA-A*02:01 HLA-B*07:02 HLA-B*35:01	0.11 0.62 0.77	0.6962	Inside
27	88	96	VALESPLAL	9	HLA-B*07:02 HLA-B*35:01	0.73 0.3	0.2992	
28	89	97	ALESPLALL	9	HLA-A*02:01	0.28	-0.0293	
29	102	110	VIMNAITSL	9	HLA-A*02:01	0.21	-0.0440	

ISSN No:-2456-2165

					HLA-B*07:02	0.69		
30	144	152	DTSEVTSFY	9	HLA-B*35:01	0.17	0.2645	
31	152	160	YPSAFQEHL	9	HLA-B*07:02	0.17	-0.2505	
					HLA-B*35:01	0.05		
32	305	313	GSFIDGRVW	9	HLA-B*57:01	0.03	-0.7686	
33	332	340	GRYVIYKRY	9	HLA-B*27:05	0.04	-0.2096	
34	412	420	FLYQRGSSY	9	HLA-B*35:01	0.12	0.2572	
					HLA-B*07:02	0.43	0.1122	
35	479	487	HRNHTLRGV	9	HLA-B*27:05	0.54	0.1784	
36	484	492	LRGVFGTML	9	HLA-B*27:05	0.88	-0.0092	
37	548	456	ISNTLFGEF	9	HLA-B*57:01	0.41	0.2587	

Supplementary Table 11 : TCD8^{+} cell epitope predict in this study

N°	Start	End	Length	Peptide	HLA Alleles	Percentile Rank	VAXIJEN 2.0 score	TMHMM result
1	86	100	15	KQVALESPLALLSTE	HLA-DRB1*03:01	8.70	**0.4735**	outside
2	221	235	15	FSTLRSINLDDNQNR	HLA-DRB1*03:01	7.80	**0.6869**	outside
3	278	292	15	DGQYHEKDLDVITLF	HLA-DRB3*01:01	5.70	**0.7537**	outside
4	279	293	15	GQYHEKDLDVITLFK	HLA-DRB3*01:01	5.70	**0.7371**	outside
5	376	390	15	IKVSTSLSEDPVLTV	HLA-DRB3*01:01	6.50	**0.4333**	outside
6	377	391	15	KVSTSLSEDPVLTVS	HLA-DRB3*01:01	6.60	**0.4406**	outside
7	378	392	15	VSTSLSEDPVLTVSP	HLA-DRB3*01:01 HLA-DRB1*03:01	7.30 8.50	**0.6321**	outside
8	379	393	15	STSLSEDPVLTVSPN	HLA-DRB3*01:01 HLA-DRB1*03:01	7.90 8.50	**0.6043**	outside
9	463	477	15	NPCVTGVYTDPYPLV	HLA-DRB3*01:01	2.80	**0.8689**	outside
10	464	478	15	PCVTGVYTDPYPLVF	HLA-DRB3*01:01	2.70	**1.0015**	outside
11	465	479	15	CVTGVYTDPYPLVFH	HLA-DRB3*01:01	2.70	**0.7999**	outside
12	466	480	15	VTGVYTDPYPLVFHR	HLA-DRB3*01:01	2.70	**0.7708**	outside
13	467	481	15	TGVYTDPYPLVFHRN	HLA-DRB3*01:01	2.70	**0.7173**	outside
14	468	482	15	GVYTDPYPLVFHRNH	HLA-DRB3*01:01	9.00	**0.8078**	outside
15	98	112	15	STESVIMNAITSLSY	HLA-DRB1*03:01	7.70	**0.4159**	Inside
16	99	113	15	TESVIMNAITSLSYQ	HLA-DRB1*03:01	7.60	**0.5179**	Inside

ISSN No:-2456-2165

17	207	221	15	ALGVLRTSATGRVFF	HLA-DRB1*03:01	5.70	**0.4429**	Inside
18	222	236	15	STLRSINLDDNQNRK	HLA-DRB1*03:01	4.60	**0.8267**	Inside
19	223	237	15	TLRSINLDDNQNRKS	HLA-DRB1*03:01	3.60	**1.0223**	Inside
20	224	238	15	LRSINLDDNQNRKSC	HLA-DRB1*03:01	3.40	**1.1027**	Inside
21	225	239	15	RSINLDDNQNRKSCS	HLA-DRB1*03:01	4.20	**0.9260**	Inside
22	226	240	15	SINLDDNQNRKSCSV	HLA-DRB1*03:01	7.80	**0.8244**	Inside
23	227	241	15	INLDDNQNRKSCSVS	HLA-DRB1*03:01	7.80	**0.8787**	Inside
24	280	294	15	QYHEKDLDVITLFKD	HLA-DRB3*01:01	5.60	**0.4709**	Inside
25	83	97	15	RVYKQVALESPLALL	HLA-DRB1*03:01	9.60	**-0.0480**	
26	84	98	15	VYKQVALESPLALLS	HLA-DRB1*03:01	8.70	**0.1430**	
27	85	99	15	YKQVALESPLALLST	HLA-DRB1*03:01	8.60	**0.3832**	
28	100	114	15	ESVIMNAITSLSYQI	HLA-DRB1*03:01	7.50	**0.3989**	
29	207	221	15	AIGVLRTSATGRVLF	HLA-DRB1*03:01	5.70	**-0.0532**	
30	281	295	15	YHEKDLDVITLFKDW	HLA-DRB3*01:01	5.50	**0.3893**	
31	282	296	15	HEKDLDVITLFKDWV	HLA-DRB3*01:01	5.50	**0.0915**	
32	285	299	15	DLDVITLFKDWVANY	HLA-DRB3*01:01 HLA-DRB1*03:01	4.40 4.10	**-0.3744**	
33	286	300	15	LDVITLFKDWVANYP	HLA-DRB3*01:01 HLA-DRB1*03:01	4.30 4.30	**-0.5611**	
34	287	301	15	DVITLFKDWVANYPG	HLA-DRB3*01:01 HLA-DRB1*03:01	5.20 4.50	**-0.8508**	
35	288	302	15	VITLFKDWVANYPGV	HLA-DRB3*01:01 HLA-DRB1*03:01	4.60 4.30	**-0.6161**	

Supplementary Table 12 : TDC4^{+} cell epitope predict in this study