Design and Evaluation of Multi-Epitope Based Vaccine against Newcastle Disease Virus

Muhammad Hassan Farooq
00000274632
MS Industrial Biotechnology

Supervisor:
Dr. Najam us Sahar Sadaf Zaidi

Atta-ur-Rahman School of Applied Biosciences (ASAB)
National University of Sciences & Technology (NUST)

Design and Evaluation of Multi-Epitope Based Vaccine against Newcastle Disease Virus

By

Muhammad Hassan Farooq
00000274632

A thesis submitted to partial fulfillment of the requirement for the degree of Masters in Sciences.

In

Industrial Biotechnology

Atta-ur-Rahman School of Applied Biosciences (ASAB)
National University of Sciences & Technology (NUST)


بسم الله الرحمن الرحيم

THESIS ACCEPTANCE CERTIFICATE

It is certified that the final copy of MS thesis written by **Mr. Muhammad Hassan Farooq,** registration No. **00000274632**, of ASAB has been vetted by undersigned, found complete in all aspects as per NUST status/regulations, is free of plagiarism, errors and mistakes and is accepted as partial fulfilment for the award of MS/MPhil degree. It is further certified that necessary amendments as pointed by GEC members of the scholar has also been incorporated in the said thesis.

Signature: _____

Name of Supervisor: **Dr. Najam us Sahar Sadaf Zaidi**

Date: _____

Signature: _____

HOD Industrial Biotechnology: **Dr. Saadia Andleeb**

Date: _____

Signature: _____

Principal ASAB: **Dr. Hussnain A. Janjua**

Date: _____

CERTIFICATE OF PLAGIARISM

It is confirmed that MS thesis entitled **"Design and Evaluation of Multi-Epitope Based Vaccine against Newcastle Disease Virus"** of **Mr. Muhammad Hassan Farooq,** registration no. **00000274632** has been me. I undertake that:

1. The thesis has significant new work as compared to already reported one. No figure, table, sentence or section has been copied from previous work except when placed under quotation marks and duly referenced.

2. The presented work is original and author's own work as no data fabrication was found in the results i.e. there is no plagiarism.

3. This thesis has been checked for Plagiarism and the Turnitin report endorsed by Supervisor is attached.

Signature of Supervisor
Dr. Najam us Sahar Sadaf Zaidi
Tenured Associate Professor
Atta-ur-Rahman School of Applied Biosciences (ASAB)
National University of Sciences and Technology (NUST)
Islamabad, Pakistan

DECLERATION

I certify that this research work titled **"Design and Evaluation of Multi-Epitope Based Vaccine against Newcastle Disease Virus"** is my own work. The work has not been presented elsewhere for assessment. The material that has been used from other sources has been properly acknowledged/referred.

_____

Muhammad Hassan Farooq

…this thesis is dedicated to my mother

and my supervisor…

ACKNOWLEDGEMENTS

All praises to Almighty ALLAH the most merciful, who is the Lord of this universe, Creator of mankind and king of wisdom. He gifted us with the ability to observe, think and research by providing us dynamic sense. All respect and honors to **the Holy Prophet Hazrat Muhammad (Peace be upon Him).** The eternal educator, the everlasting source of leadership, the torchbearer of truth and messenger to the whole world.

It gives us immense pleasure and satisfaction to write in favor of those peoples who played very important role in completion of this research and contributed to it in one way or another. I feel highly privileged and honor to express special thanks to my worthy supervisor **Dr. Najam us Sahar Sadaf Zaidi** for providing me opportunity to work under her supervision and fulfilling all the requirements and lab facilities needed for my research.

I would also like to thankful to **Dr. Amjad Ali**, for his incredible help at the time of need. He was always ready to help me and support me. My deepest gratitude is for GEC members; **Dr. Saadia Andleeb** and **Dr. Muhammad Tahir**, without their support and concern this would have not been possible.

I am thankful to HoD Industrial Biotechnology **Dr. Saadia Andleeb** and Principal ASAB **Dr. Hussnain A. Janjua** for their incredible support and motivation.

Bundles of thanks and prayers for my senior lab fellows Saleha & Azikah and my research mates Manahil, Saleha, Afra, Zomia, Shbana and Kashaf for their help and wonderful company during this research at ASAB.

No language in this world has these words that explain the gratitude, love which supports efforts of my parents in each and every stage of education specifically and my entire life. Thanks to my friends Noor, Zunaira, Humaira, Maryam, Huzaifa, Abu Bakar, Abdullah, Mudassir, Dawood and Osama to be a source of constant support.

Table of Contents

Chapter 1

Introduction 1

Objectives 7

Chapter 2

Literature Review 8

2.1. Newcastle Disease Virus 9

2.2. Target Viral Proteins 10

2.2.1. Haemagglutinin-Neuraminidase 10

2.2.2. Fusion Protein 12

2.3. Currently Available Vaccines 13

2.4. Limitations of Current Live Attenuated Vaccines 14

2.5. Multi-epitope Vaccines Designed via In-Silico Approaches 16

2.6. Adjuvants in Multi-epitope Vaccine 18

2.7. Linkers in Multi-epitope Vaccine 20

2.8. Multi-epitope Vaccine against NDV 20

Chapter 3

Materials and Methods 22

3.1. Selection of Target Proteins 23

3.2. Proteome Retrieval 23

3.2.1. Selection of HN Protein Sequences 23

3.2.2. Selection of F Protein Sequences 24

3.3. Development of Consensus Sequence 28

3.4. Prediction of B cell Epitopes 28

3 5. Prediction of T cell Epitopes 29

3 6. Selection of Overlapping B cell and T cell Epitopes 29

3.7. Prediction of IFN-γ Epitopes 29

3.8. Evaluation of Pan-Proteomic Epitopes 30

3.8.1. Non- Homologous 30

3.8.2. Immunogenic Potential 30

3.8.3. Antigenic Potential 30

3.8.4. Allergen Prediction 30

3.8.5. Conservancy 31

3.9. Multi-epitope Vaccine Design 31

3.9.1. Epitope-Epitope Interactions and Arrangement 31

3.9.2. Addition of Adjuvant 31

3.10. Physiochemical Properties of the Vaccine Construct 32

3.11. 3D Structure modeling of the Vaccine Construct 32

3.12. Refinement of 3D Modeled Vaccine Construct 32

3.13. Validation of 3D Modeled Vaccine Construct 33

3.13.1. Ramachandran Plot Analysis 33

3.13.2. ProSA-web-Z Score Evaluation 33

3.13.3. Verify3D 33

3.14. Molecular Docking of the Vaccine with Toll like Receptors 34

3.15. Docking Interactions 34

Chapter 4

Results 35

4.1. Proteome Retrieval and Consensus Sequence 36

4.2. Predicted B cell Epitopes 38

4.2.1. Prediction of B cell Epitopes from HN Protein 38

4.2.2. Prediction of B cell Epitopes from F Protein 40

4.3. Predicted T cell Epitopes 43

4.3.1. Prediction of T cell Epitopes from HN Protein 43

4.3.2. Prediction of T cell Epitopes from F Protein 45

4.4. Overlapping Epitopes 47

4.4.1. Overlapping Epitopes of HN Protein 47

4.4.2. Overlapping Epitopes of F Protein 48

4.5 Evaluation of Epitopes 49

4.5.1. Immunogenic Potential 49

4.5.2. Antigenic Potential 50

4.5.3. Allergenecity Prediction 54

4.5.4. Conservance 56

4.6. Predicted IFN-Gamma Inducing Epitopes 57

4.7. Finalized Epitopes 58

4.8. Multi-Epitope Vaccine Design 59

4.8.1. Epitope-Epitope Interactions and Arrangement 59

4.8.2. Addition of Adjuvant 60

4.9. Physiochemical Properties of Vaccine Construct 61

4.10. Modeling of the 3D Structure of Vaccine Construct 62

4.11. Refinement and Validation of 3D Modeled Vaccine Construct 63

4.12. Validation of 3D structure 64

4.12.1. Ramachandran Plot Analysis 64

4.12.2. ProSA-web-Z Score Evaluation 65

4.12.3. Verify3D 65

4.13. Molecular Docking of the vaccine with Toll like Receptors 67

4.13.1 Molecular Docking of Vaccine Construct with TLR4 68

4.13.2 Molecular Docking of Vaccine Construct with TLR2 type 1 68

4.13.3 Molecular Docking of Vaccine Construct with TLR2 type 2 69

4.14 Intersurface statistics of vaccine construct with Toll like Structures 69

4.14. Docking Interactions 71

Chapter 5

Discussion 73

Chapter 6

Conclusion 74

Future Perspectives 75

Chapter 7

References 76

List of Tables

Table 2.1: NDV Strains used in production of live vaccine. 14

Table 2.2 : Limitations of live attenuated vaccines and their solutions. 16

Table 2.3: Type of adjuvants and their respective immune response. 18

Table 3.1: Genbank IDs of all selected HN protein sequences. 23

Table 3.2: Genbank IDs of selected F protein sequences. 24

Table 4.1: Consensus sequence of HN and F Protein . 36

Table 4.2: B cell epitopes predicted by ABCpred from HN consensus sequence 38

Table 4.3: B cell epitopes predicted by SVMtrip from HN consensus sequence 40

Table 4.4: B cell epitopes predicted by ABCpred from F consensus sequence 40

Table 4.5: B cell epitopes predicted by SVMtrip from F consensus sequence 42

Table 4.6: T cell epitopes predicted by HLApred from HN consensus sequence 43

Table 4.7: T cell epitopes predicted by HLApred from F consensus sequence 45

Table 4.8: Overlapping epitopes of HN protein. 47

Table 4.9: Overlapping epitopes of F protein. 48

Table 4.10: HN protein epitopes with positive immunogenicity value 49

Table 4.11: F protein epitopes with positive immunogenicity value. 51

Table 4.12: HN protein epitopes with antigenic potential 52

Table 4.13: F protein epitopes with antigenic potential 53

Table 4.14: HN protein epitopes with non-allergen nature 54

Table 4.15: HN protein epitopes with non-allergen nature 55

Table 4.16: Conservancy of HN protein epitopes 56

Table 4.17: Conservancy of F protein epitopes 56

Table 4.18: IFN gamma inducing HN protein epitopes 57

Table 4.19: IFN gamma inducing F protein epitopes 57

Table 4.20: Properties of finalized epitopes in HN Protein 58

Table 4.21: Properties of finalized epitopes in F Protein 58

Table 4.22: Epitope-Epitope interactions and their arrangements 59

List of Figures

Figure 2.1: Schematic diagram of NDV genome encoding six proteins; P (Phosphoprotein), NP (Nucleocapsid protein), F (Fusion protein), Matrix protein (M), Heamagglutinin-Neuraminidase (HN) and Large protein (L). 9

Figure 2.2: The structure of Newcastle disease virus. 10

Figure 2.3: Schematic diagram of F protein activation. Inactive precursor F0 is cleaved into F1 and F2 subunits by host proteolytic enzymes. 12

Figure 2.4: Overview of an overall immune response after antigen exposure 17

Figure 4.1: Development of consensus sequence of (a) HN Protein (b) F Protein through UGene tool 37

Figure 4.2: (a) Schematic diagram of proposed vaccine construct. **(b)** Overall sequence of the construct**. (c)** Color key for the vaccine construct 61

Figure 4.3: 3D structure of the vaccine construct made by ITasser. 62

Figure 4.4: (a) Structure and properties of the construct before refining **(b)** Structure and properties after refining. 63

Figure 4.5: Ramachandran plot analysis. 64

Figure 4.6: ProSA-web-Z score analysis. 65

Figure 4.7: Verify 3D verification. 65

Figure 4.8(a): Interaction of Vaccine construct with TLR4. 68

Figure 4.8(b): Interaction of Vaccine construct with TLR2 type 1. 68

Figure 4.8(c): Interaction of Vaccine construct with TLR2 type 2. 69

Figure 4.9(a): Inter-surface analysis between construct and TLR4 69

Figure 4.9(b): Inter-surface analysis between construct and TLR2 type 1 ..... **70**

Figure 4.9(c): Inter-surface analysis between construct and TLR2 type 2 71

Figure 4.10: Protein-protein interaction between vaccine construct between;
(a) TLR4 **(b)** TLR2 type 1 **(c)** TLR2 type 2 72

List of Abbreviations

AAvV	Avian Avulavirus
APMV	Avian Paramyxovirus
BLAST	Basic Local Alignment Search Tool
BVDV	Bovine Viral Diarrhea Virus
CTB	Cholera Toxin B
ELISA	Enzyme Linked Immunosorbent Assay
END	Exotic Newcastle Disease
F	Fusion Protein
HADDOCK	High Ambiguity Driven protein-protein DOCKing
HN	Haemagglutinin-Neuraminidase
IEDB	Immune Epitope Database
IFN- γ	Interferon Gamma
IL	Interleukin
M	Matrix Protein
MDT	Mean Death Time
MHC	Membrane Histocompatibility Complex

NCBI	National Center for Biotechnology Information
NDV	Newcastle Disease Virus
NK	Natural Killer Cells
NP	Nucleocapsid Protein
OIE	World Organization for Animal Health
PAMP	Pathogen Associated Molecular Patterns
PRR	Pattern Recognition Receptors
PVC	Potential Vaccine Candidates
RSV	Respiratory Syncytial Virus
TLR	Toll Like Receptors
TNF-β	Tumor Necrosis Factor-Beta

ABSTRACT

Abstract

Newcastle disease virus is a single stranded RNA virus of negative polarity, is the main causative agent of NDV infection in chickens. Despite the availability of commercial live attenuated vaccine, the outbreaks are common in poultry industry causing huge economic losses. The main factors responsible for the vaccine ineffectiveness are; high mutation rate, genetic variability of the virus and other factors causing immunosuppression in birds. In this study, computational and immune-informatics approach was used to design a multi- epitope vaccine based on the sequences reported from Pakistan. The 8 prioritized epitopes were predicted from structural proteins; HN and F of NDV. These non-host homologous epitopes were capable of inducing strong B cell, T cell and IFN- γ response against the virus. Furthermore, the epitopes were arranged on the basis of epitope - epitope interaction analysis and linked with each other through GPGPG linkers. An adjuvant (CTB) was added at the N terminal to enhance the immunogenicity of the construct. Construct was then modeled, refined and evaluated using online tools. Significant docking score signal out towards great interaction between vaccine construct and TLR receptors, thus enabling the vaccine to induce TLR activation which will be followed by an amplified immune response against the virus. These results show that the proposed vaccine construct can induce a strong innate and adaptive immune response against NDV in chickens, however, experimental validation will be necessary to confirm its potential. In future, this study can be utilized in finding genetic diversity among NDV strains reported globally from different regions of the world and can also be helpful in detection of genetic determinants associated vaccine resistance in vaccinated chickens.

INTRODUCTION

Introduction

Newcastle disease virus is an enveloped virus containing single stranded negative sense RNA genome. It belongs to the genus *Avulavirus* of *Paramyxoviridae* family (Peeters & Koch, 2019). The virus is zoonotic in nature and causative agent of a highly acute and contagious respiratory disease in birds called Newcastle disease. The virus was first discovered in the year 1926 at a place called Newcastle, United Kingdom and since then it has been causing fatal effects on the poultry industries resulting in huge losses in economy (Phale, 2018). In addition to chickens, the NDV virus can also cause infrequent diseases in other wild birds as well. Initially, the Newcastle disease virus was taxonomically named as Avian paramyxovirus 1 (APMV-1), however, the name have been recently changed to Avian Avulavirus 1 (AAvV-1). The genus Avulavirus consists of 19 Avula viruses (AAvV-1 to AAvV-19) including the Newcastle disease virus (AAvV-1). The genus Avulavirus comes under the family Paramyxoviridae of the order Mononegavirales (Amarasinghe et al., 2017).

The virus is pleomorphic, with a single strand of RNA genome consisting of 15,186 bases of nucleotides and helical capsid symmetry. Furthermore, it contains six transcriptional units which codes for the following six proteins - three envelope and three core proteins. One of the envelope protein exhibits the activities of both haemagglutinin and neuraminidase (HN), the second one exhibits the fusion activity (F) and the third one (M) is located inside the envelope. The core proteins are nucleocapsid protein (NP), the large protein (L) and phosphoprotein (P) (Czeglédi et al., 2006). HN and F protein are the immunogenic proteins in nature, also the most significant proteins in the determining the virulence and infectivity of the virus because of their role in activating membrane fusion and viral entry in host cells. The order of the above mentioned six proteins is as; 3'-NP-PM-F-HN-L-5' (Phale, 2018). Just like the other members of the family paramyxoviridae, NDV also forms some additional proteins V and W, which actually arise as a result of translation of P protein by the alternative mRNA. The alternative mRNAs are produced by the RNA editing during the transcription of P gene (Rao et al., 2020).

NDV has been classified into two further classes; Class I and class II. The Class II comprises of 16 genotypes. The genotypes III-IX and XI-XVI from Class II are all

virulent in nature. (Diel et al., 2013; Courtney et al., 2013). Different strains attack different organs and the intensity of the pathogenicity also varies from strain to strain. NDV have also been categorized in the following pathotypes, based on the type and intensity of the symptoms they cause in chickens. Three pathogenic strains are classified as; the low virulent lentogenic strains, moderately virulent mesogenic strains and highly virulent velogenic strains (Beard & Hanson 1984). The velogenic strains are further classified into two categories; lethal hemorrhage causing viscerotropic-velogenic strains and neurological disorders causing neurotropic-velogenic strains (Alexander, 1988).

On the basis of different basic amino acids present at proteolytic fusion cleavage site of F protein, the virulent strains of NDV are determined. This aspect can really prove molecular basis of variation in virulence. The presence of basic amino acids at fusion site is responsible for mediation of fusion and cell-cell and cell-virus interactions. Phenylalanine (F) at position 117 and three or more residues of either lysine (K) or arginine (R) at position 113 are almost found in every virulent strain of NDV (Brown, V. & Bevins, 2017).

Enveloped viruses enter the host cells through different fusion methods. It can either be (i) Direct fusion; in which the envelope membrane of the virus fuses with cell membrane of the host or (ii) receptor-mediated endocytosis; in which a receptor is involved in the fusion process and nucleocapsid is translocated inside the host cell as a result of endocytosis (Dimitrov, 2004). When virus enters the host body, then the proteases present inside the host are responsible for cleaving the precursor F0 Fusion protein into F1 and F2 proteins. This cleavage will be responsible for the initiation of infection by fusion and several other homeolytic processes. The viruses which contain an active cleavage site are virulent because any type of proteases, present inside the host body, can cleave the F protein and start the infection. Whereas, in the absence of an active cleavage site, only trypsin and trypsin mediated enzymes can cause the cleavage of the fusion proteins. As these enzymes are mostly present at respiratory and intestinal tracts, so this can result in restricted host site replication. (Huang et al, 2004; Nagai & Clenk, 1977).

NDV is a causative agent of a highly contagious disease of poultry; New Castle Disease. More than 250 species of the birds are reported to be prone to this virus, making this

disease to be one of the most deadliest disease of poultry industry. On the basis of the strains responsible for causing the infection, NDV can be divided into three forms; (a) mild or lentogenic form (b) moderate or mesogenic form (c) virulent or velogenic form. The highly virulent form of new castle disease is also called as Exotic Newcastle Disease (END). END is regarded as the most virulent disease of poultry. It affects all species of birds. This disease is so virulent that most of the birds die without displaying any clinical symptoms. Luckily, END has no known human health effects (Brown & Torres, 2008; Wakamatsu et al, 2006). The highly pathogenic form of Newcastle disease has been listed in World Organization for Animal Health (OIE) Terrestrial Animal Health Code and must be reported to the OIE (OIE 2004).

The most common mode of transmission of NDV is through direct contact with carrier or diseased birds. The virus is shed into their feces by infected birds thus contaminating the area surrounding them. Then the transmission can easily occur by aerosol exposure or by coming direct in-contact with either the infected bird itself or with the oral and fecal discharge, contaminated food, equipment, water and even by clothing. It can sustain for many weeks in the environment especially in winter (Estola & Hovi, 1979).

The signs and symptoms of the infection varies on the basis of type of NDV strain causing the infection. It can also depend upon species, health and age of the host bird. The range of incubation period can be from four to six days, once the virus enters inside the host body. Signs and symptoms can appear within young birds of 2-12 days (Average 5). The spread by aerosol exposure is faster as compared to the spread by fecal-oral route. Respiratory symptoms which predominate in NDV infection involves; sneezing, coughing, rales and gasping. Nervous signs of paralyzed legs and wings, tremors, clonic spasms, circling and complete paralysis can also be seen (Absalón et al., 2019). Nervous signs are mostly observed in exotic birds and cormants. Diarrhea accompanied by nervous symptoms is a frequent symptom observed in pigeons. Either partial or complete termination of egg production can be observed. Eggs will feel irregular in shape, color or size and can contain transparent watery albumen (Roberts et al., 2011). In well vaccinated birds, the signs and symptoms are not properly visible except for the decline in egg production. However, they can still shed virus in saliva and feces and can easily spread

infection to poorly vaccinated birds, which may be able to develop torticollis, ataxia and other symptoms 2-14 days after the exposure and can recover with supportive care. In case of viscerotropic velogenic strain, the gross leisons can be seen. The histopathology changes which mainly observed are; pneumonia, tracheitis, myocarditis, pericarditis, atropy of Bursa Fabricious, nephritis interstitial, encephalitis and splenitis (Etriwati et al., 2017).

Newcastle disease has a high mortality rate in domestic birds and is responsible for the devastating economic losses in poultry industry (Lancaster,1976; Spardbrow et al., 1988). Since most of the birds are not routinely vaccinated, the disease can easily spread with direct contact between diseased and healthy birds thus resulting in mass mortality of poultry chickens. Although there have been major advancements in treatment, diagnosis and vaccination against Newcastle disease since 1950s, it is still considered as a major threat to the poultry industry ultimately a menace to the economy. (Phale, 2018). USA faced a loss of US 162 million dollars during the last major outbreak of Newcastle disease in California during 2002-03. This outbreak alone was responsible for mortality of 4 million birds (Cattoli et al., 2011).

In Pakistan, the poultry industry is the second largest industry after textile industry. The poultry sector is one of the most important zones of agriculture industry in Pakistan with a massive contribution of 1.3% in national GDP. (Rehan et al., 2019). The recent reported ND outbreak in Pakistan during 2012 caused a loss of 6 billion Pakistani Rupees. Since 1960s, the poultry industry has been commercially providing a significant contribution of 26.8 % meat production and 5.76% of eggs to the population. (Hussain et al., 2015). Poultry sector in Pakistan saw 20-30% growth per anum in early 1970s and about 10-15% in 1980s making it one of the most important sector contributing in economy with a consumption of almost 4% per anum (Sadiq, 2004). The poultry industry is also an employment source for almost more than 1.5 million people (GOP, 2016).

Despite the availability and extensive use of live attenuated vaccine against NDV, significant outbreaks have been observed all over the world. The commercial live attenuated vaccine contains the LeSota strain of NDV. Master seed immunogenicity

tests are used to evaluate the immunogenicity of vaccine. This vaccine is capable of inducing immunity even with very mild reactions. It further contains penicillin, streptomycin and fungizone as bacteriostatic and fungistatic agents respectively. This vaccine is recommended via drinking water or aerosol spray for vaccination against the Newcastle disease in chickens. Aerosol vaccination is only recommended in case of revaccination. The vaccine is stored at 35-45°F (2-7°C) without freezing it (Cornax et al., 2012; Su et al., 2018).

Some of the most common reasons which are responsible for outbreaks even with administration of vaccine are; the unequal and uneven mass administration of the vaccine in large commercial settings, administering the same vaccine in multi-age birds which are roaming freely in different areas possibly carrying and transferring the virus (Dimitrov et al., 2017). Ineffectiveness of already existing vaccine can also be due to high mutation rates in RNA viruses, intra-species variation of NDV strains based on geographical location and other factors which can cause immunosuppression in birds. All these factors can really affect the effectiveness of traditional live attenuated vaccine. Moreover, the procedure to develop it is very tedious and costly (Absalón et al., 2019).

Advancements in bioinformatics and computational biology have caused a greater level of understanding about vaccine and its design. With the help of these tools, not only the pathogenic antigens can be predicted, but their antigenicity and immunogenicity can also be evaluated. Through reverse vaccinology, potential vaccine targets (PVCs) can easily be identified through genome mining by using computer aided tools (Sette & Rappuoli 2010).

Objective of the study:

The objectives of this study included;

- Selection of HN and F Protein sequences reported from Pakistan and finding their consensus sequence.
- Prediction of B cell and T cell overlapping epitopes and their characterization.
- Designing of a multi-epitope vaccine, its 3D structure evaluation and docking analysis with TLRs.

The development of vaccine will help in eradication of virus and its frequent outbreaks. Moreover, it will be a successful alternative for the commercial live attenuated vaccine with its limitations.

LITERATURE REVIEW

2.1. Newcastle Disease Virus

Newcastle disease is a fatal disease of poultry with detrimental effects on the economy. The virus responsible for causing this lethal disease is Newcastle disease virus. It was previously named as avian paramyxovirus type-1 (AMPV-1). However, the name was changed to avian avulavirus after some taxonomical changes. In this study, the name avulavirus and Newcastle disease virus will be used interchangeably. NDV is a zoonotic virus and can be further distinguished into further pathotypes on the basis of MDT (mean death time) in chicken embryos. Lentogenic, being the non-virulent ones, have MDT of more than 90 hours. Mesogenic, the moderately virulent ones, have MDT between 60-90 hours. The most virulent velogenic strains have MDT of less than 60hours (Brown & Bervins, 2017).

The virion of the viruses in paramyxoviridae varies in their size (150-300nm) and can take spherical, pleomorphic or the filamentous shape. The virus contains single stranded , non-segmented, negative sense RNA genome of approximately 15kb in size. The genome encodes the six genes; P (Phosphoprotein), NP (Nucleocapsid protein), F (Fusion protein), Matrix protein (M), Heamagglutinin-Neuraminidase (HN) and Large protein (L). The order of the above mentioned six proteins is as; 3'-NP-P-M-F-HN-L-5' (as shown in the figure 1) (L′vov et al., 2015).

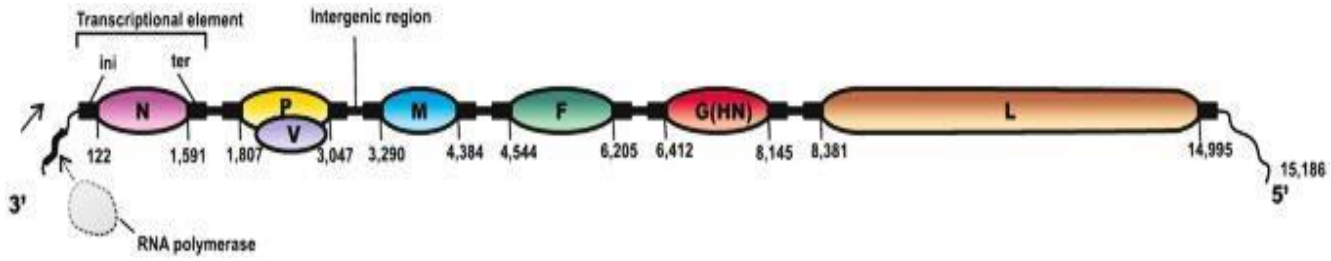


Figure 2.1: Schematic diagram of NDV genome encoding six proteins; P (Phosphoprotein), NP (Nucleocapsid protein), F (Fusion protein), Matrix protein (M), Heamagglutinin-Neuraminidase (HN) and Large protein (L) (L′vov et al., 2015).

HN and F proteins are anchored in the envelope membrane of NDV and are responsible for viral attachment with the cellular receptors. The F protein can specifically take part in the fusion process which enables the entry of viral particle into the cell by enabling

fusion of viral envelope membrane and host cell membrane. Moreover, these proteins are also responsible for the viral release from the host cells. Both of these proteins are immunogenic in nature, and determine the virulence and infectivity of the virus. Therefore, in this study, these two proteins were selected as target proteins for epitope prediction in constructing a polyepitope vaccine.

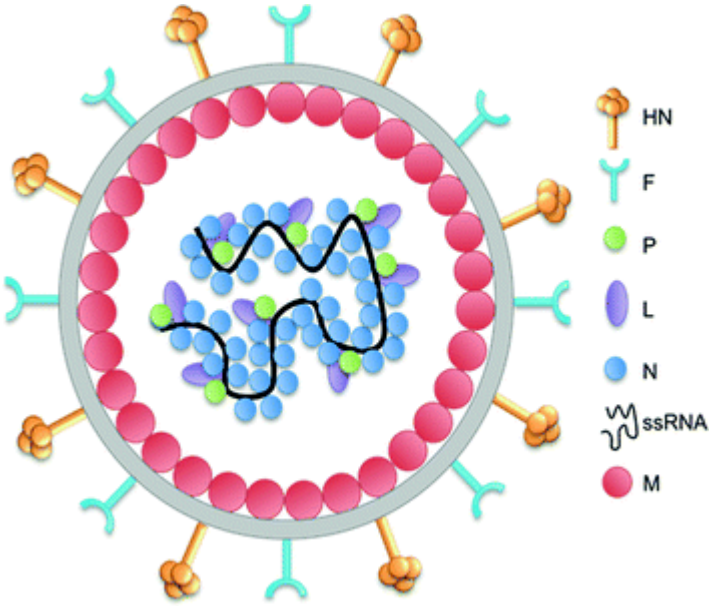


Figure 2.2: The structure of Newcastle disease virus (Thomas & Walmsley, 2018).

2.2. Target Viral Proteins

2.2.1. Haemagglutinin-Neuraminidase

Haemagglutinin – Neuraminidase is a vital antigenic determinant of NDV, with a length of about 1998 nucleotides that encodes for a long polypeptide consisting of 577 amino acids (Phale, 2018). The HN can easily bind with sialic acid, so it enables the virus to bind with those receptors containing sialic acid. After binding to the sialic acid containing receptor, it mediates the neuraminidase activity i.e. enzymatic cleavage of the sialic acid. Along with these activities, it also aids in fusion activity by interacting it with F protein. HN protein has always been considered as important immunogenic protein for vaccine development.

In 1992, Juan McEwen and his co-workers developed a recombinant vaccine, which involved expressing an epitope of HN protein of influenza virus in flagellin of *Salmonella*. Thus, this synthetic recombinant vaccine was successful in evaluating the potential of HN epitope as a potential vaccine (McEwen et al., 1992).

In 2003, Hua Li designed a recombinant protein which contained neutralizing epitopes of heamagglutinin of H392 influenza virus. These epitopes were combined and the recombinant immunogen was expressed in E.coli. ELISA results confirmed a higher titer of antibodies was induced against the H392 influenza virus as a result of using neutralizing epitope of HN (Li, 2003).

In 2008, Pallavi Somvanshi, Vijay Singh, and P.K. Seth used haemmaglutinin and neuraminidase proteins of Influenza virus (Strain: H5N1) for prediction of epitopes. These epitopes were further proved beneficial in vaccine development against the above mentioned virus (Somvanshi et al., 2008).

In 2009, Wilfred Ndifon and his co-workers worked on the design of influenza vaccine after determining the differential neutralization efficiency of the HN epitopes. This also helped them in understanding the amino acid changes in HA can affect in increase of mutated virus neutralization. This occurs as HA mutation can result in steric interference among the antibodies (Ndifon et al., 2009).

In 2014, Florian Krammer along with Peter Palese and John Steel worked on influenza virus and its vaccine strategies on the basis of conserved regions of Haemagglutinin and Neuraminidase (Krammer et al., 2014).

In 2019, Mahmudul Hasan and his co-workers used the approach of reverse vaccinology for development of a subunit multi-epitope vaccine. This polyepitope vaccine was designed against avian influenza A (H7N9) virus using haemagglutinin and matrix protein 1 as the target proteins for prediction of epitopes. Both these proteins were regarded as the most antigenic ones thus further confirming the HN immunogenic potential and its role in development of subunit vaccines (Hasan et al., 2019).

2.2.2. Fusion Protein

F protein or fusion protein has a length of about 1792 nucleotides that encodes for a long polypeptide chain consisting of 553 amino acids (Chambers et al., 1986). The basic role of F protein is to mediate viral entry. This is achieved by fusion of viral envelope membrane and host plasma membrane. Fusion can be done either directly or via receptor mediated endocytosis. The F protein is synthesized in an inactive precursor form called F0. Host enzymes are required for its cleavage. As a result of the cleavage by host cell proteolytic enzymes, F1 and F2 subunits are formed which are joined together by a disulfide bridge as shown in the figure (Dutch, 2010).

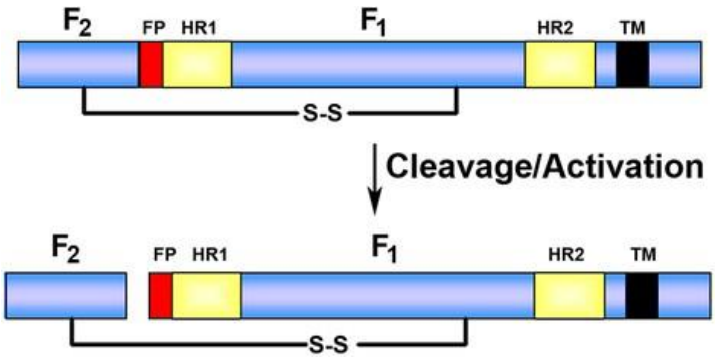


Figure 2.3: Schematic diagram of F protein activation. Inactive precursor F0 is cleaved into F1 and F2 subunits by host proteolytic enzymes (Dutch, 2010).

F protein is considered as a major antigenic determinant and has been used in designing vaccines and developing ELISA kits, for detection and seroprevalance, due to its antigenic and immunogenic nature. The F protein along with HN was considered as the target viral proteins, in this study, due to the above mentioned properties. In 2012, Patricia Sastre and her co-workers developed recombinant fusion protein based ELISA for the seroprevalence of human metanemumovirus and RSV (respiratory syncytial virus). Thus, evaluating the potential of F protein as a major antigenic determinant and can be highly preferred as target protein in predicting T-cell and B-cell epitopes while designing a multi-epitope vaccine (Sastre et al., 2012).

Emad A. Hashish and his co-workers, in 2013, used a multiepitope fusion antigen against the BVDV (bovine viral diarrhea virus) and ETEC (enterotoxigenic *E.coli*). The mutiepitope induced neutralizing antibodies against both pathogens (Hashish et al., 2013).

In 2019, Gaafar and her co-workers designed a multiepitope vaccine against Peste des Petits Ruminants Virus or small ruminant morbillivirus. The epitopes were predicted from the fusion (F), haemagglutinin (H), matrix (M) and nucleocapsid (N) proteins of the virus using different insilico tools and softwares (Gaafar et al., 2019).

In 2020, Muhammad Tahir ul Qamar and his team used the approach of reverse vaccinology in designing a multi epitope vaccine. The vaccine was designed against RSV (respiratory syncytial virus) by using its fusion and glycoprotein as the target proteins for prediction of T-cell and B-cell epitopes. These epitopes ultimately showed strong interactions with human TLRs. Hence, a multiepitope construct was designed using computational and bioinformatics tools against RSV (Tahir Ul Qamar et al., 2020).

2.3. Currently Available Vaccines

Currently, there are many live vaccines available against NDV in the market. Apart from the velogenic strains of NDV, there are eight different strains which are being used in live vaccines against the Newcastle disease virus (Table 2.1).

The thermostable vaccines against NDV show a major resistance against elevated temperatures rather than the heat labile ones. Different strains of Newcastle disease virus varies a lot in thermostability. When a vaccine prepared from a thermostable strain, it can retain its potential even outside the cold chain storage for specific amount of time. The two basic procedures for developing a thermostable vaccine against NDV is as follows; either isolate the naturally occurring variants of the virus which are thermostable in nature or thernostability of the variant can be increased in the laboratory through artificial selection. The antigenicity and the thermostability are two of the important key factors for an adequate vaccine (Grimes, 2002).

Table 2.1: NDV Strains used in production of live vaccine (Grimes, 2002).

Strains	Pathotype	Description
F	Lentogenic	Usually used in young chickens but suitable for use as a vaccine in chickens of all ages.
B1	Lentogenic	Slightly more virulent than F, used as a vaccine in chickens of all ages.
La Sota	Lentogenic	Often causes post vaccination respiratory signs, used as a booster vaccine in flocks vaccinated with F or B1.
V4	Avirulent	Used in chickens of all ages.
V4-HR	Avirulent	Heat Resistant V4, thermostable, used in chickens of all ages.
I-2	Avirulent	Thermostable, used in chickens of all ages.
Mukteswar	Mesogenic	An invasive strain, used as a booster vaccine. Can cause adverse reactions (respiratory distress, loss of weight or drop in egg production and even death) if used in partially immune chickens. Usually administered by injection.
Komarov	Mesogenic	Less pathogenic than Mukteswar, used as booster vaccine. Usually administered by injection.

2.4. Limitations of Current Live Attenuated Vaccines

Despite the availability and extensive use of live attenuated vaccine against NDV, significant outbreaks have been observed all over the world. Therefore, NDV is still a major threat for the poultry industry and responsible for huge economic losses. There can be certain factors, which can be responsible for the limitation of the commercial vaccine currently available in the market. One of the most important factors, which are responsible for the outbreaks despite the vaccination is; antigenic differences between the vaccine itself and the strains which caused the outbreak. In such cases, vaccine is unable to control the viral replication and its spread. Therefore, there always is a need of new vaccine which genotype matches the outbreak strain (Liu et al., 2017).

As we know that the live attenuated vaccine is unable to successfully vaccinate all the birds of the poultry flock, the virus which is being shed from the vaccinated bird can infect the non-vaccinated or even the poorly vaccinated birds. The shedding of virus from

the use of live attenuated vaccine is major concern and one of the most prominent limitations. Hence new strategies are required to tackle the problem of virus shedding.

Another important limitation in use of live attenuated vaccine is the presence of maternal bodies in chickens up to age of three weeks. These maternal antibodies can obstruct the replication and infectivity of the live attenuated viruses. When these maternal antibodies are weaned, the viruses in vaccination can cause growth retardation or respiratory diseases (Dortmans et al., 2011). This problem could've been easily prevented if the live attenuated vaccine was not the only available option for vaccination against NDV in a commercial setting.

Some of the most common reasons which are responsible for outbreaks even with administration of vaccine are; the unequal and uneven mass administration of the vaccine in large commercial settings, administering the same vaccine in multi-age birds which are roaming freely in different areas possibly carrying and transferring the virus (Dimitrov et al., 2017).

Other reason responsible for the limitation of the already available commercial vaccine can include high mutation rates of the virus. As the RNA genome does not have any proof reading mechanism during its replication, therefore, the mutation rate in RNA replication is much higher than the DNA replication, which actually has the proof reading mechanism (Manoharan et al., 2018). Some other reasons may include intra-species variation of NDV strains based on geographical location and immunosuppression in birds. All these factors can really affect the effectiveness of traditional live attenuated vaccine.

Table 2.2 depicts all the possible limitations which are faced by the live attenuated vaccines along with their possible solutions. As we can see that one of the major solution of these limitations is to design a synthetic vaccine which is actually the main objective of this study. Moreover, the process of developing the traditional live attenuated vaccines is not cost and time efficient. Advancements in bioinformatics and computational biology have enabled a greater level of understanding about vaccine design. With the help of these tools, not only the antigens of a pathogen can be predicted easily, but their

immunogenicity can also be evaluated. Through reverse vaccinology, potential vaccine targets (PVCs) can easily be identified through genome mining by using computer aided tools.

Table 2.2 : Limitations of live attenuated vaccines and their solutions (Soema et al., 2015).

Limitations	Potential Solutions
Dependence on egg-based production	• Cell culture-based production of virus • Recombinant antigens • Synthetic Vaccines
Regulatory approval procedures	• Mock-up vaccines to generate regulatory dossier
Limited worldwide vaccine availability	• Technology transfer of vaccine production methods • Dose spraying by the addition of adjuvants or alternative administration routes • Increase stability and shelf life of vaccines to prevent vaccine loss in unfavorable conditions

2.5. Multi-epitope Vaccines Designed via In-Silico Approaches

Multi-epitope vaccine is that type of vaccine which is designed by joining different epitopes together to form a vaccine construct. These epitopes are predicted through online tools within the proteins of that specific pathogen against which the vaccine is being made. These vaccines contain epitopes for activation of T cells (Cytotoxic T cells and Helper T cells) and B cells (Azmi et al., 2014).

When the antigenic epitopes are exposed to the B cells; the B cells, with the help of surface B cell receptors, can recognize the antigen. In addition to this, B cells also require interactions between T_h cells and B cells via co stimulators for their activation and hence inducing a proper immune response. The T cells are activated through the antigen presentation on MHC (membrane histocompatibility complex). The antigens are presented after being processed through exogenous or endogenous pathways depending upon the type of antigen. Once the antigen is processed and presented on MHC molecules, the T_h cells can easily recognize the processed antigen presented on the MHC

via T-cell receptor (TCR) present on the surface of T_h cells. Once the Th1 cells are activated, they further activate the CTLs (Cytotoxic T cells) by secreting specific cytokines. So, it can be said that the activation of CTLs is dependent on cytokines secreted by activated TH1 cells. These examples of such specific cytokines are interferon-gamma (IFN-γ), interleukin-2 (IL-2) and tumor necrosis factor-beta (TNF-β) (Moyle and Toth, 2013). These cytokines further induce the antigen presenting cells (APCs) to secrete some co-stimulatory molecules responsible for activation of CTLs. In addition to that, these cytokines are also responsible for the activation and proliferation of natural killer cells (NK) along with Th1 and Th2 cells. These Th2 cells can further cause the proliferation of B cells along with the production of second wave of cytokines, interleukins 4,5,10 and 14, hence strengthens the humoral response as well. Therefore, multi-epitope vaccine contains both T-cell and B-cell epitopes which can induce a strong humoral and cell mediated response (Lei et al., 2019).

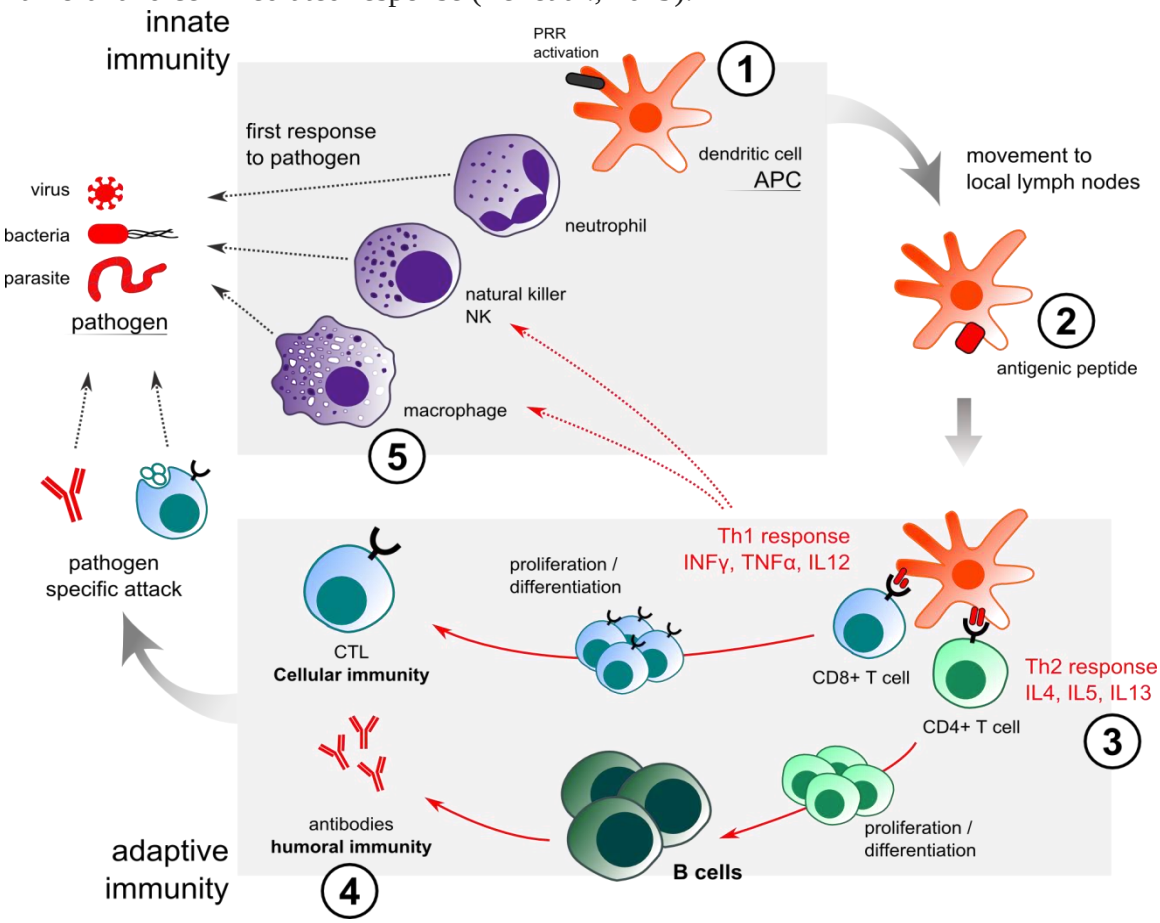


Figure 2.4: Overview of an overall immune response after exposure to an antigen (Nimbalkar et al., 2018).

2.6. Adjuvants in Multi-epitope Vaccine

An adjuvant is an agent which helps in enhancing the immune response of a vaccine. The adjuvant attached to the vaccine helps in stimulation of PAMPs (pathogen associated molecular patterns). These PAMPs are recognized by the PRRs (pattern recognition receptors) which are present of host cells. The pathogen is identified through its respective PAMPs. During the construction of a multi-epitope vaccine, the adjuvants are added to enhance the PAMPs present in the epitopes used. As we know that multi-epitope vaccine only contains the epitopes of the virus, therefore the adjuvants are sometimes necessary for the identification of the virus or pathogen so that a proper immune response can be generated against it (Apostólico et al., 2016).

Adjuvants can be classified on the basis of their physiochemical properties, source, administration route and their mechanism of action. Sometimes the adjuvants are also classified depending upon the type of immune response they can generate. Some adjuvants are capable of generating Th1 immune response while others can generate the Th2 (Mehrabi et al., 2018).

Table 2.3: Type of adjuvants and their respective immune response (Mehrabi et al., 2018).

Immune Response	Antigens	Adjuvants	References
Humoral immunity /rarely induce cellular immune responses	Diphtheria, tetanus, pertussis (DTP), *Haemophilus influenza type b*, pneumococcal conjugates, hepatitis A , B and polio virus	Aluminium phosphate or hydroxide	(Kool et al., 2012)
High levels of IgG/ does not increase IgE production	Diphtheria, tetanus, pertussis vaccines	Calcium phosphate	. (Jiang et al., 2004)
Humoral and cellular immunity	Human papillomavirus (HPV)(1)	Adjuvant System 04 (AS04) (consists of aluminum hydroxide and monophosphoryl lipid A [MPL])	(Schwarz and Leo, 2008)
Humoral and cellular immunity	*Edwardsiella tarda*, veterinary vaccines	Freund's incomplete adjuvant (FIA)	(Jiao et al., 2010)

Cellular immunity	Malaria, HIV and cancer vaccine trials	Montanide	(Reed et al., 2009)
Eliciting both humoral and cellular immune responses	Pandemic flu (GSK)	Squalene	Leroux-Roels et al., 2007)
Cellular immunity	New Castle Disease virus (NDV), infectious bronchitis virus (IBV)	Oil Based Emulsions (e.g., MF59)	(Jansen et al., 2006; Kool et al., 2012)
Cellular and humoral immunity	*M. tuberculosis*, Tetanus toxoid (TT), Diphtheria toxoid (DT)	Liposomes	(Mishra et al., 2007)
Induced a mixed Th1/Th2 response	Influenza virus antigens	ISCOMs	(Sjölander et al., 2001)
Unique ability to stimulate cell-mediated immunity and to enhance antibody production	Veterinary vaccines	Saponins	(Reed et al., 2009; Sun et al., 2009)
Enhance cellular immune responses through a variety of mechanisms/ humoral immunity	Plasmid DNA	Cytokines	(Zheng et al., 2014)
Cellular immunity	Against hepatitis B virus core antigen	PLGA	(Chong et al., 2005)
(Ataman-Önal et al., 2006)	Cellular and humoral immunity	HIV-1 p24 protein	PLA
(Cooper et al., 2005)	Cell-mediated immune and humoral responses	Hepatitis B virus, ovalbumin	CpG-motifs (CpG-ODNs)

Cellular / humoral immunity	*H. pylori* antigens	Bacterial lipopeptide, Lipopolysaccharide, Bacterial toxins (CT, LT)	(Marchetti et al., 1998; Reed et al., 2009)
Enhanced mucosal and humoral immunity / cellular immunity	HBsAg – Snake and scorpion venom	Chitosan	(Mohammadpourdounighi et al., 2010; Wen et al., 2011; Mohammadpour Dounighi et al., 2012; Farhadian et al., 2015; Mohammadpour Dounighi et al., 2016)

2.7. Linkers in Multi-epitope Vaccine

While developing a multi-epitope vaccine construct, linkers play an important role in joining the epitopes together. Moreover, the attachment of an adjuvant with the epitopes is also done with the help of linkers. The selection of linkers for the construction of vaccine construct is important because linkers can directly influence the orientation and stability of the construct. Although different linkers vary in length and composition but can highly affect the overall structure (George and Heringa, 2002). Some of the most commonly used linkers, which are extensively studied and used in literature, are as follows; AAY, GPGPG, EAAAK, KK, GG, GGS and SSL etc. The likers containing the glycine molecules are preferred because of their small size, stability, flexibility, high solubility and their resistance towards the proteolytic reactions (Kavoosi et al., 2007).

2.8. Multi-epitope Vaccine against NDV

In this study, a multi epitope vaccine has been designed through integrating the approaches of reverse vaccinology, proteomics and immunoinformatics. The proteomics approach was used to find a consensus sequence of HN and F protein, from sequences reported all over in Pakistan. The prioritized vaccine epitopes were determined through reverse vaccinology, and it was important that the prioritized epitopes were non-chick homologs and extracellular and secretory in nature (Kumar

Jaiswal et al., 2017; Hasan et al., 2019). Using immunoimformatics approach, the structure of vaccine was modeled and refined. At last, the vaccine construct was docked with TLR2-1, TLR2-2 and TLR4 of *Gallus gallus* (chicken) to evaluate the successful interactions between the receptors and designed vaccine. Thus, the designed vaccine will induce immune responses specifically against the pathogen by containing conserved regions which can generate B-cell (CD4) and T-cell (CD8) responses(Setter & Fikes, 2003; Tu et al., 2014).

MATERIALS AND METHODS

3.1. Selection of Target Proteins

The viral structural proteins; heamagglutinin-neuraminidase (HN) and fusion protein (F) were selected as target viral proteins for epitope prediction. Both these proteins are envelope proteins of Newcastle disease virus. HN and F protein are the immunogenic proteins in nature, also the most significant proteins in the determining the virulence and infectivity of the virus because of their role in activating membrane fusion and viral entry in host cells (Phale, 2018). These PAMP (pathogen associated molecular patterns) containing epitopes from these surface proteins can directly be recognized via host PRRs (pattern recognition receptors) and hence a strong immune response can be generated (Mogenson, 2009).

3.2. Proteome Retrieval

The NCBI (National Center for Biotechnology Information) database (https://www.ncbi.nlm.nih.gov/) was used be used to retrieve the protein sequences. NCBI is a national database which provides genomic and biomedical information (O'Leary et al., 2016). A total of 27 HN protein sequences and 100 F protein sequences were retrieved from NCBI. The details of all the selected sequences of both HN and F proteins are given in the table 3.1 and 3.2 respectively. Only those sequences with complete cds were selected. All these sequences are reported from different cities of Pakistan.

3.2.1 Selection of HN Protein Sequences

Table 3.1: Genbank IDs of all selected HN protein sequences.

Sr No.	GenBank ID	DBSOURCE	City/District
1.	AEX55097.1	JN682207.1	Islamabad
2.	AEX55096.1	JN682206.1	Islamabad
3.	AEX55095.1	JN682205.1	Rawalpindi
4.	AEX55094.1	JN682204.1	Attock
5.	AEX55093.1	JN682203.1	Rawalpindi
6.	AEX55092.1	JN682202.1	Islamabad
7.	AEX55091.1	JN682201.1	Rawalpindi
8.	AEX55090.1	JN682200.1	Lahore

9.	AGK41181.1	JX436339.1	Lahore
10.	QCF28566.1	MH891654.1	Lahore
11.	QCF28565.1	MH891653.1	Lahore
12.	ALL27113.1	KR676409.1	Attock
13.	ALL27112.1	KR676408.1	Wazirabad
14.	ALL27124.1	KR676420.1	Gujranwala
15.	ALL27123.1	KR676419.1	Lahore
16.	ALL27122.1	KR676418.1	Lahore
17.	ALL27121.1	KR676417.1	Lahore
18.	ALL27120.1	KR676416.1	Sheikhupura
19.	ALL27119.1	KR676415.1	Gujranwala
20.	ALL27118.1	KR676414.1	Multan
21.	ALL27117.1	KR676413.1	Attock
22.	ALL27116.1	KR676412.1	Lahore
23.	ALL27115.1	KR676411.1	Lahore
24.	ALL27114.1	KR676410.1	Lahore
25.	ALL27111.1	KR676407.1	Narowal
26.	ALL27110.1	KR676406.1	Peshawar
27.	ALL27109.1	KR676405.1	Kohat

3.2.2 Selection of F Protein Sequences

Table 3.2: Genbank IDs of selected F protein sequences.

Sr No.	GenBank ID	DBSOURCE	City/District
1.	AFI81995.1	JQ517285.1	Lahore
2.	AXY66660.1	MG686609.1	Lahore
3.	AXY66659.1	MG686608.1	Lahore
4.	AXY66658.1	MG686607.1	Lahore
5.	AXY66657.1	MG686606.1	Sialkot
6.	AXY66656.1	MG686605.1	Sialkot

7.	AXY66655.1	MG686604.1	Lahore
8.	AXY66654.1	MG686603.1	Karachi
9.	AXY66653.1	MG686602.1	Lahore
10.	AXY66652.1	MG686601.1	Lahore
11.	AXY66651.1	MG686600.1	Gujranwala
12.	AXY66650.1	MG686599.1	Gujranwala
13.	AXY66649.1	MG686598.1	Gujranwala
14.	AXY66648.1	MG686597.1	Gujranwala
15.	AXY66647.1	MG686596.1	Sheikhupura
16.	AXY66646.1	MG686595.1	Sheikhupura
17.	AXY66645.1	MG686594.1	Sheikhupura
18.	AXY66644.1	MG686593.1	Sialkot
19.	AXY66643.1	MG686592.1	Islamabad
20.	AXY66642.1	MG686591.1	Gujranwala
21.	AXY66641.1	MG686590.1	Gujranwala
22.	AXY66640.1	MG686589.1	Sheikhupura
23.	AXY66639.1	MG686588.1	Sialkot
24.	AXY66638.1	MG686587.1	Sialkot
25.	AXY66637.1	MG686586.1	Islamabad
26.	AXY66636.1	MG686585.1	Sialkot
27.	AXY66635.1	MG686584.1	Lahore
28.	AXY66634.1	MG686583.1	Peshawar
29.	AXY66633.1	MG686582.1	Lahore
30.	AXY66632.1	MG686581.1	Karachi
31.	AXF73534.1	MH120424.1	Faisalabad
32.	ALL27108.1	KR676404.1	Gujranwala
33.	ALL27107.1	KR676403.1	Lahore
34.	ALL27106.1	KR676402.1	Lahore
35.	ALL27105.1	KR676401.1	Lahore
36.	ALL27104.1	KR676400.1	Sheikhupura

37.	ALL27103.1	KR676399.1	Gujranwala
38.	ALL27102.1	KR676398.1	Multan
39.	ALL27101.1	KR676397.1	Attock
40.	ALL27100.1	KR676396.1	Lahore
41.	ALL27099.1	KR676395.1	Lahore
42.	ALL27098.1	KR676394.1	Lahore
43.	ALL27097.1	KR676393.1	Attock
44.	ALL27096.1	KR676392.1	Wazirabad
45.	ALL27095.1	KR676391.1	Narowal
46.	ALL27094.1	KR676390.1	Peshawar
47.	ALL27093.1	KR676389.1	Kohat
48.	AGK41177.1	JX436341.1	Sheikhupura
49.	QCE30397.1	MH891148.1	Azad Jammu Kahmir
50.	QCE30396.1	MH891147.1	Azad Jammu Kahmir
51.	AGK41180.1	JX436344.1	Lahore
52.	AGK41179.1	JX436343.1	Okara
53.	AGK41178.1	JX436342.1	Gujranwala
54.	AGK41176.1	JX436340.1	Faisalabad
55.	AOM52883.1	KU862299.1	Lahore
56.	AOM52882.1	KU862298.1	Lahore
57.	AOM52881.1	KU862297.1	Lahore
58.	AOM52880.1	KU862296.1	Lahore
59.	AOM52879.1	KU862295.1	Karachi
60.	AOM52878.1	KU862294.1	Lahore
61.	AOM52877.1	KU862293.1	Karachi
62.	AOM52876.1	KU862292.1	Kamoki
63.	AOM52875.1	KU862291.1	Patoki
64.	AOM52874.1	KU862290.1	Lahore
65.	AOM52873.1	KU862289.1	Lahore
66.	AOM52872.1	KU862288.1	Lahore

67.	AOM52871.1	KU862287.1	Lahore
68.	AOM52870.1	KU862286.1	Lahore
69.	AOM52869.1	KU862285.1	Lahore
70.	AOM52868.1	KU862284.1	Lahore
71.	AOM52867.1	KU862283.1	Lahore
72.	QCF28564.1	MH891652.1	Lahore
73.	QCF28563.1	MH891651.1	Lahore
74.	AEX55081.1	JN682191.1	Islamabad
75.	AEX55080.1	JN682190.1	Islamabad
76.	AEX55079.1	JN682189.1	Rawalpindi
77.	AEX55078.1	JN682188.1	Attock
78.	AEX55077.1	JN682187.1	Rawalpindi
79	AEX55076.1	JN682186.1	Islamabad
80.	AEX55075.1	JN682185.1	Rawalpindi
81.	AEX55074.1	JN682184.1	Lahore
82.	QCX35388.1	MK006017.1	Kasur
83.	QCX35387.1	MK006016.1	Banka Cheema
84.	QCX35386.1	MK006015.1	Banka Cheema
85.	QCX35385.1	MK006014.1	Banka Cheema
86.	QCX35384.1	MK006013.1	Gharoo
87.	QCX35383.1	MK006012.1	Kasur
88.	QCX35382.1	MK006011.1	Mirpur Khas
89.	QCX35381.1	MK006010.1	Sargodha
90.	QCX35380.1	MK006009.1	Sargodha
91.	AMR55429.1	KU644588.1	Lahore
92.	AMR55427.1	KU644586.1	Lahore
93.	AWU46622.1	MH392224.1	Karachi
94.	AWU46620.1	MH392223.1	Karachi
95.	AWU46618.1	MH392222.1	Karachi
96.	ARE67979.2	KY076043.2	Narang Mandi

97.	ARE67978.1	KY076039.2	Gujranwala
98.	ARE67977.1	KY076038.2	Bhai Phairu
99.	ARE67973.1	KY076034.2	Chak Shehzad
100.	ARE67969.1	KY076030.2	Sheikhupura

3.3. Development of Consensus Sequence

The consensus proteome sequences of both proteins were obtained using **UGENE** software tool (http://ugene.net/). UGENE is a software which is used for bioinformatics approaches. This tool helps users in viewing, analyzing and annotating biological data; in form of NSG assemblies, multiple sequence alignment and phylogenetic trees etc (Okonechnikov et al., 2012).

All 27 selected sequences of HN protein and 100 sequences of F protein were combined in a separate notepad and were multiply aligned using UGENE software. After multiple alignments of all those sequences, consensus sequence was obtained. This consensus sequences were further used in prediction of B cell, T cell and IFN epitopes.

3.4. Prediction of B cell Epitopes

The consensus sequence obtained from UGENE software was further used to predict the B cell epitopes. The online tools which were used for his purpose were the **SVMtrip** (http://sysbio.unl.edu/SVMTriP/) and **ABCpred** (http://crdd.osdd.net/raghava/abcpred/). ABCpred is an online tool which predicts the linear continuous epitope sequences of B cells (Saha and Raghava, 2006). The basic purpose of SVMtrip is also to predict the B cell epitopes from the given query sequence (Yao et al., 2012).

The query sequences of HN and F protein were uploaded separately on both of these softwares,. The threshold level set for both of these online prediction tools was above 0.8 for SVMtrip and 0.51 for ABCpred. Only those epitopes with values above the threshold level were selected.

3.5. Prediction of T cell Epitopes

The consensus sequence obtained from UGENE software was also further used to predict the T cell epitopes. The online prediction tools which were used for this purpose were **HLApred** (http://crdd.osdd.net/raghava/hlapred/ref.html). HLA pred is an online tool which predicts the HLA binding reigons in the given antigenic query sequence. The HLA binding regions for both class I and class II can be selected (Brusic *et al*, 1994).

The query sequences were uploaded on the online HLApred server. Both HLA classes were selected for the prediction of binding reigons within our query sequence. 3 % threshold level was selected. Epitopes showing high score and binding affinities were selected.

3.6. Selection of Overlapping B and T cell Epitopes

B cell epitopes and T cell epitopes predicted through their respective online prediction were taken into consideration and only those were selected which overlapped with each other. In other words, only those B cell epitopes were proceeded further which had T cell epitope sequences present within them. So, in this way a single epitope is capable of generating both B cell and T cell response.

These overlapping epitopes were further scrutinized on the on the basis of their ability to produce Interferon gamma, which was predicted using IFNepitope online server.

3.7. Prediction of IFN-γ Epitopes

The online tool used for the prediction of IFN-γ epitopes was **IFNepitope** (http://crdd.osdd.net/raghava/ifnepitope/) online server. This online prediction tool is specifically designed for the prediction of IFN-γ inducing epitopes (Dhanda et al., 2013).

The previously prioritized epitopes were uploaded on the IFNepitope server and only those epitopes were further selected which showed the ability to induce IFN-γ response. Hence, B cell epitopes overlapping with predicted T cell and IFN- γ were prioritized and separated and regarded as pan-proteomic epitopes.

3.8. Evaluation of Pan-Proteomic Epitopes

3.8.1. Non- Homologous

These separated epitopes were further evaluated with **Blastp** (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins) against chicken proteome (Altschul et al., 1990; Gish & States, 1993). As we are designing a vaccine for poultry chickens, therefore the vaccine should not contain any epitopes or sequences showing homology with the chicken proteome. So, by doing blastp the chances of having epitopes homologous to chicken proteins can be avoided.

3.8.2. Immunogenic Potential

The immunogenic potential of the epitopes was evaluated by **MHC1 immunogenicity** score provided by Immune Epitope Database **(IEDB)** server (http://tools.iedb.org/immunogenicity/) (Vita et al., 2019). The previously selected epitopes were uploaded on the server and their immunogenic potential was determined in this way. All such epitopes which showed negative immunogenic values were discarded.

3.8.3. Antigenic Potential

The Antigenicity potential was evaluated by using **VaxiJen v2.0** (http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html) and epitopes above the threshold value (0.5) (Doytchinova et al., 2007). All those epitope below the threshold level and regarded non-antigen by the vaxijen tool were immediately discarded.

3.8.4 Allergen Prediction

AllergenFP (https://ddg-pharmfac.net/AllergenFP/) and **Allertop** (https://www.ddg-pharmfac.net/AllerTOP/) were the online tools which were used to determine allergenecity prediction of the epitopes (Dimitrov et al.,2013). Only those epitopes showing non allergen nature were selected. Hence, the epitopes which were finalized showed positive immunogenicity score, high antigenicity score and were non allergen in nature.

3.8.5 Conservancy

Blastp (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins) was used to find the conservancy of the selected epitopes among the all NDV strains (Altschul et al., 1990; Gish & States, 1993). The epitopes with maximum conservancy were selected and proceeded further.

These finalized epitopes possessed the following characteristics; (a) B cell epitopes overlapping with predicted T cell and IFN- γ epitopes. (b) High immunogenic and antigenic values and non-allergen in nature. (c) Non Homologous to the chicken proteins and highly conserved among the NDV strains.

3.9. Multi-epitope Vaccine Design

3.9.1. Epitope-Epitope Interactions and Arrangement

To design a vaccine construct comprising of epitopes, it is necessary to determine the arrangement of these epitopes. In order to determine the arrangement, the epitopes were evaluated for their binding affinities with each other using online HADDOCK (https://wenmr.science.uu.nl/enmr/services/HADDOCK2.2/) server (Dominguez et al., 2003).

For this purpose, the 3D structure of these epitopes was predicted by using I-TASSER (https://zhanglab.ccmb.med.umich.edu/I-TASSER/) (Roy et al., 2010). Initially; each epitope was evaluated for its binding energy with each other epitope. The pair of epitope showing the maximum energy was selected and then joined with flexible GPGPG linker. This initial pair was then further evaluated with the remaining epitopes for the evaluation of maximum binding energy. In this way the arrangement pattern of epitopes was determined for the vaccine design.

3.9.2. Addition of Adjuvant

An immunogenic adjuvant; Cholera cytotoxin B (CTB) (WP_000593522.1) was additionally added in the N terminal of the prioritized epitopes through EAAK linker. The finalized vaccine design consists of epitopes, linked together with GPGPG linkers, and an adjuvant which was linked to the epitopes via EAAK linker (Dar et al., 2019).

3.10. Physiochemical Properties of the Vaccine Construct

ProtParam (https://web.expasy.org/protparam/) online server was used to calculate the physiochemical properties of the vaccine construct (Gasteiger et al., 2005). The properties which are as follows; molecular weight, instability index and aliphatic index are very crucial to determine the nature of the vaccine and its stability. The FASTA sequence is uploaded on the online ProtParam server and it predicts all the above mentioned properties.

3.11. 3D Structure Modeling of the Vaccine Construct

The three dimensional modeling of vaccine construct was done by ITASSER (https://zhanglab.ccmb.med.umich.edu/I-TASSER/) server ((Roy et al., 2010). ITASSER is an online tool used for structure homology and prediction of 3D structures of the proteins.

The FASTA sequence of the finalized vaccine construct was uploaded on the online ITASSER server. The server predicted 5 models with different confidence scores. The confidences of the predicted models were estimated by C-score. Basically, C-score is calculated based on the significance of threading template alignments and the convergence parameters of the structure assembly simulations. 3-D model with highest value of C score was selected because high value of C-score signifies the model has high confidence (Ikram et al., 2018).

3.12. Refinement of 3D Modeled Vaccine Construct

Galaxy refine (http://galaxy.seoklab.org/cgi-bin/submit.cgi?type=REFINE) is an online tool of bioinformatics which can provide many services including the prediction of the protein structure, the refinement of the structure it can also predict the interaction between different proteins (Ko et al.,2012; Heo et al., 2013; HeeShin et al., 2014) .

The selected model, predicted and modeled by ITASSER was further processed by online Galaxy Refine server to obtain a more refined model of the vaccine construct with many residues in favorable regions.

3.13. 3D Validation of 3D Modeled Vaccine Construct

For the validation of the structure; Ramachandran plot analysis on RAMPAGE server, Prosa web Z score and Verify 3D were used.

3.13.1. Ramachandran Plot Analysis

The most commonly used tool for the validation of any 3D structure of any protein or vaccine construct is done by ramachandran plot analysis. Ramachandran plot has torsion angles of the amino acids in any given protein. The torsion angles involved in this plot are; psi (ψ) and phi (φ). So, by ramachandran plot, we can estimate which torsion angles are possible and permitted. Through this plot, we can know a lot about the protein structure (Ramachandran et al., 1968).

The structure of our vaccine construct, in pdb format, was uploaded on RAMPAGE (http://mordred.bioc.cam.ac.uk/~rapper/rampage.php) server and the ramchandran analysis was obtained (Wang et al., 2016).

3.13.2. ProSA-web-Z Score Evaluation

Z score is basically used to evaluate the structure of protein structures. From all the other alternatives present for the query structure, the software recognizes the native fold and gives score on the basis of this. The energy of the lattice protein model and of the experimental protein is compared. The plot is formed and the energy of the query model is compared with the energy of all the protein models already available in the protein database (Satyanarayana et al., 2018).

The structure of the query vaccine construct was uploaded at **ProSA-web** (https://prosa.services.came.sbg.ac.at/prosa.php) server and the plot showing Z score was obtained (Wiederstein & Sippl (2007).

3.13.3. Verify3D

The 3D model of the vaccine construct was also evaluated through the online tool **Verify3D**. This tool uses the atomic co-ordinates of the amino acids present in the

structure. Then the coordinates of the query sequence is compared with already predicted good models and the result is shown (Eisenberg et al., 1997).

The structure of the query vaccine construct was uploaded at **Verify3D** (https://servicesn.mbi.ucla.edu/Verify3D/) server and the results were obtained (Tran et al., 2015).

3.14. Molecular Docking of the Vaccine with Toll like Receptors

Docking of vaccine structure with toll like receptors of chicken (*Gallus gallus*) was done by using **HADDOCK** (High Ambiguity Driven protein-protein DOCKing) (https://wenmr.science.uu.nl/enmr/services/HADDOCK2.2/) server (Dominguez et al., 2003).

At first the active residues of both the vaccine construct and the TLRs was obtained by **cport** (https://alcazar.science.uu.nl/services/CPORT/) server (De Vries & Bonvin, 2011). These active residues will be uploaded on the HADDOCK server along with the sequences of both ligand (vaccine construct) and receptors (TLRs).

TLR2-type 1, TLR2-type 2 and TLR4 of chicken (*Gallus gallus*) were selected to dock with the construct. The result is shown in clusters with their respective HADDOCK score. According to HADDOCK, the cluster showing the lowest HADDOCK score signifies the highest interaction between proteins.

3.15. Docking Interactions

The detailed analysis of the selected docked cluster was further done by using PDBsum (https://www.ebi.ac.uk/pdbsum/online) server (Laskowski, 2004). This gave us a detailed insight about the residues and the intermolecular forces present within the docked cluster.

RESULTS

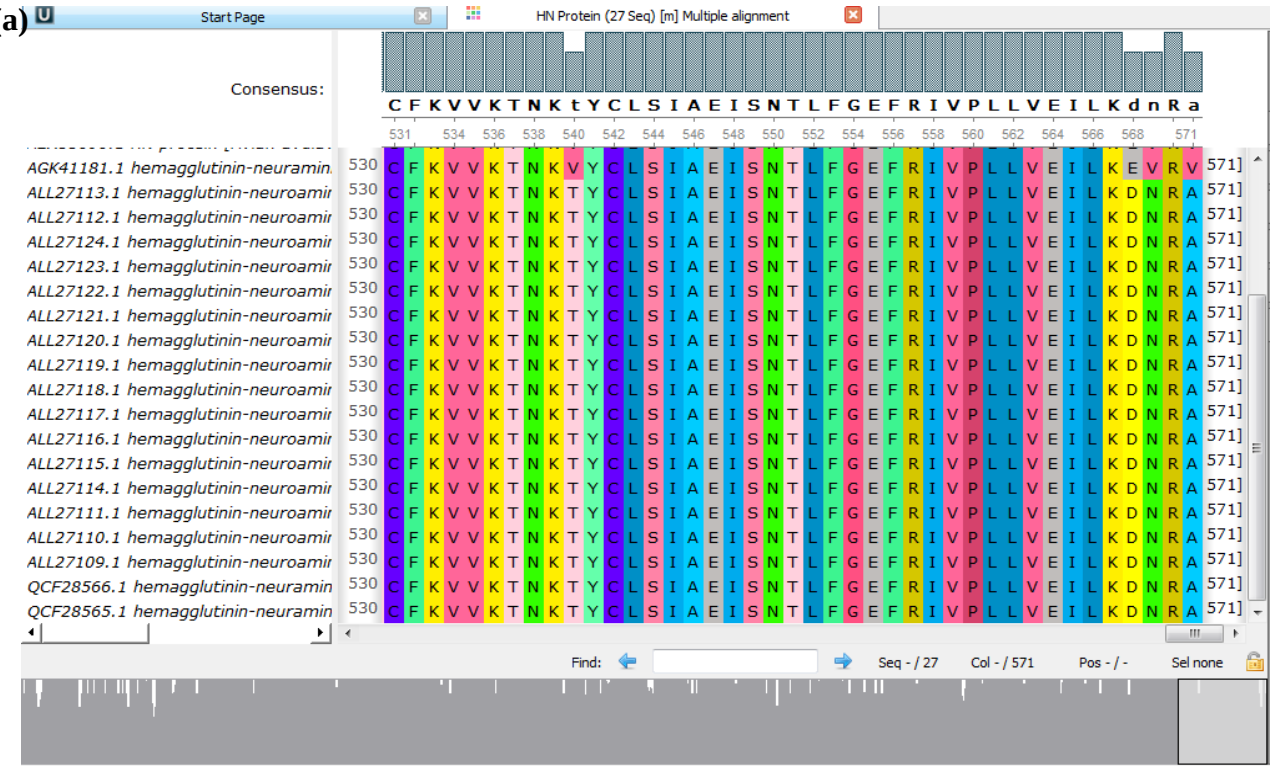
4.1. Proteome Retrieval and Development of Consensus Sequence

A total of 27 sequences of HN protein and 100 sequences of F protein were retrieved from NCBI. These selected sequences were reported from all over the Pakistan. The Sequence IDs along with their accession numbers are provided in table 4.1 for HN and F proteins. Following consensus sequences were found through UGENE software tool.

Table 4.1 : Consensus sequence of HN and F Protein.

HN Protein	MSRAVNRVMLENEEREAKNTWRLVFRIAVLLLMIMILAISAAALAYSMG TSTPRDLTGISVAISKTEDKVTSLLSSSQDVIDRIYKQVALESPLALLNTESI IMNAITSLSYQINGAANNSGCGAPVHDPDYIGGIGKELIVDDTSDVTSFYP SAYQEHLNFIPAPTTGSGCTRIPSFDMSTTHYCYTHNVILSGCRDHSHSHQ YLALGVLRTSATGKVFFSTLRSINLDDTQNRKSCSVSATPLGCDILCSKVT ETEEEDYKSVTPTSMVHGRLGFDGQYHEKDLDTTALFKDWVANYPGVG GGSFVDERVWFPVYGGLKPNSPSDTAQEGKYVIYKRYNDTCPDKQDYQI RMAKSSYKPGRFGGKRVQQAILSIKVSTSLGEDPMLTIPPNTITLMGAEGR ILTVGTSHFLYQRGSSYFSPALLYPMTISNKTATLHSPYTFNAFTRPGSVPC QASARCPNSCITGVYTDPYPLIFHRNHTLRGVFGTMLDDGQARLNPVSAV FDDISRSRVTRVSSSSTKAAYTTSTCFKVVKTNKTYCLSIAEISNTLFGEFRI VPLLVEILKDNRA
F Protein	MGSKPSTRIPVPLMLITRIMLILSYICLTSSLDGRPLAAAGIVVTGDKAVNV YTSSQTGSIIVKLLPNMPKDKEACAKAPLEAYNRTLTTLLTPLGDSIRKIQ GSVATSGGRRQKRFIGAVIGSVALGVATAAQITAAAALIQANQNAANILR LKESIAATNEAVHEVTDGLSQLSVAVGKMQQFVNDQFNNTARELDCIKIT QQVGVELNLYLTELTTVFGPQITSPALTQLTIQALYNLAGGNMDYLLTKL GVGNNQLSSLIGSGLITGYPILYDSQTQLLGIQVNLPSVGNLNNMRATYLE TLSVSTTKGFASALVPKVVTQVGSVIEELDTSYCIESDLDLYCTRIVTFPMS PGIYSCLSGNTSACMYSKTEGALTTPYMALKGSVIANCKITTCRCADPPGI ISQNYGEAVSLIDRHSCNVLSLDGITLRLSGEFDATYLKNISILDSQVIVTG NLDISTELGNVNNSISNALDKLTESNSKLDKVNVRLTSTSALITYIALTVIS LFFGVLSLGLACYLMYKQKAQQKTLLWLGNNTLDQMRATTRA

(a)



(b)

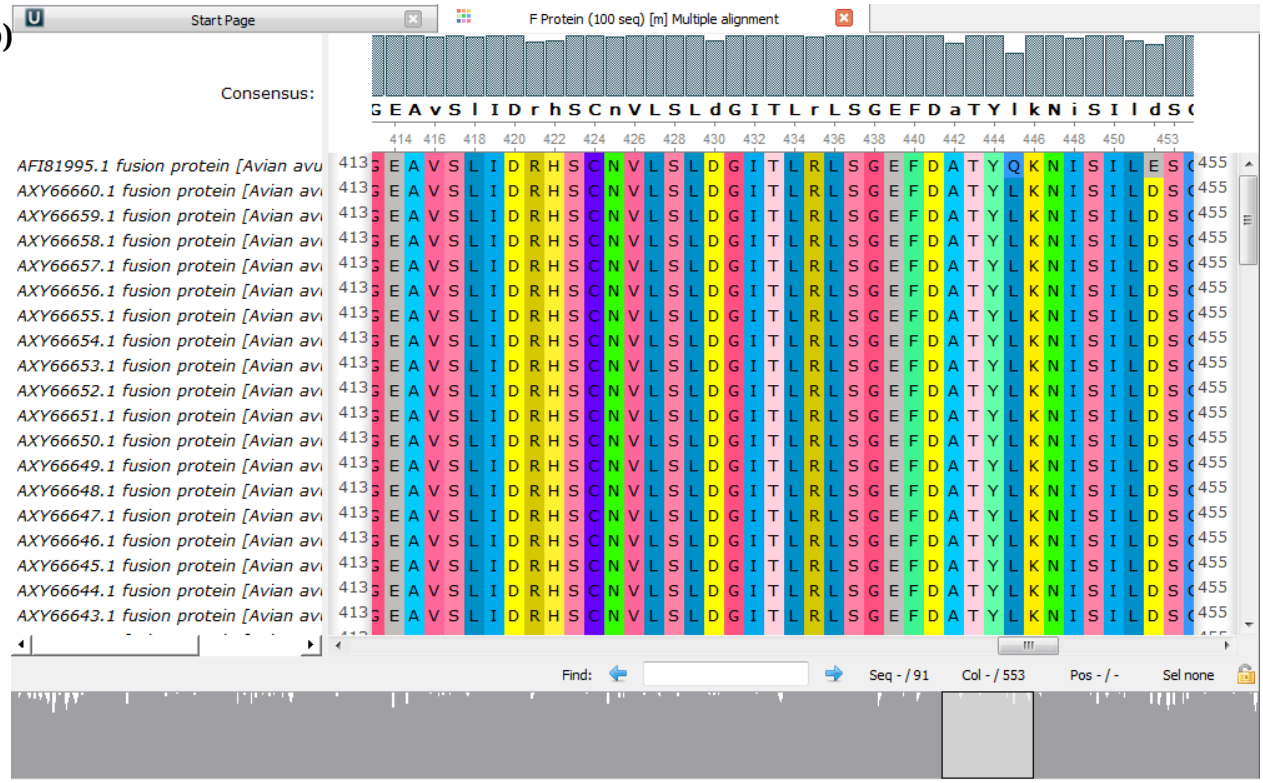


Figure 4.1:Development of consensus sequence of (a) HN Protein (b) F protein through UGene tool.

4.2. Predicted B cell Epitopes

ABCpred predicted 54 Linear B cell epitopes (Table 4.2) for HN protein and 51 B cell epitopes of F protein (Table 4.4). SVMtrip predicted 7 (Table 4.3) and 9 (Table 4.5) B cell epitopes of HN and F protein respectively.

4.2.1. Prediction of B cell Epitopes from HN Protein

Table 4.2: B cell epitopes predicted by ABCpred from HN consensus sequence. (Threshold 0.51)

Rank	Sequence	Start position	Score
1	GVFGTMLDDGQARLNP	486	0.94
2	TATLHSPYTFNAFTRP	435	0.93
3	TRPGSVPCQASARCPN	448	0.91
3	DERVWFPVYGGLKPNS	309	0.91
4	SIKVSTSLGEDPMLTI	375	0.9
5	TGVYTDPYPLIFHRNH	467	0.89
5	AKSSYKPGRFGGKRVQ	355	0.89
5	GRLGFDGQYHEKDLDT	273	0.89
6	CSKVTETEEEDYKSVT	251	0.88
6	LRSINLDDTQNRKSCS	224	0.88
6	DVTSFYPSAYQEHLNF	147	0.88
6	ESIIMNAITSLSYQIN	100	0.88
7	PDEQDYQIRMAKSSYK	345	0.87
8	HRNHTLRGVFGTMLDD	479	0.86
9	YSMGTSTPRDLTGISI	46	0.85
9	PNTITLMGAEGRILTV	392	0.85
9	SCSVSATPLGCDILCS	237	0.85
9	CGAPVHDPDYIGGIGK	123	0.85
10	KAAYTTSTCFKVVKTN	523	0.84
10	PVSAVFDDISRSRVTR	501	0.84
10	QASARCPNSCITGVYT	456	0.84
10	YQINGAANNSGCGAPV	112	0.84
11	TGSGCTRIPSFDMSTT	168	0.83

12	IAEISNTLFGEFRIVP	545	0.82
12	WVANYPGVGGGSFVDE	295	0.82
13	DTAQEGKYVIYKRYND	327	0.81
14	VKTNKTYCLSIAEISN	535	0.8
14	GKRVQQAILSIKVSTS	366	0.8
14	KELIVDDTSDVTSFYP	138	0.8
15	LSGCRDHSHSHQYLAL	193	0.79
15	MSTTHYCYTHNVILSG	180	0.79
16	QDVIDRIYKQVALESP	78	0.76
16	SYFSPALLYPMTISNK	419	0.76
16	EGRILTVGTSHFLYQR	401	0.76
16	GKVFFSTLRSINLDDT	217	0.76
17	VNRVMLENEEREAKNT	005	0.75
17	GEDPMLTIPPNTITLM	383	0.75
18	TPLGCDILCSKVTETE	243	0.74
18	PSAYQEHLNFIPAPTT	153	0.74
18	AITSLSYQINGAANNS	106	0.74
19	EFRIVPLLVEILKDNR	555	0.72
19	VTRVSSSSTKAAYTTS	514	0.72
20	PYTFNAFTRPGSVPCQ	441	0.71
20	KDLDTTALFKDWVANY	284	0.71
20	HSHSHQYLALGVLRTS	199	0.71
20	NEEREAKNTWRLVFRI	12	0.71
21	HFLYQRGSSYFSPALL	411	0.7
21	LGVLRTSATGKVFFST	208	0.7
22	IYKQVALESPLALLNT	84	0.67
22	GISIAISKTEDKVTSL	58	0.67
23	LDDGQARLNPVSAVFD	492	0.65
24	YKRYNDTCPDEQDYQI	337	0.64
25	DDISRSRVTRVSSSST	507	0.63

26	VFRIAVLLLMIMILAI	24	0.52

Table 4.3: B cell epitopes predicted by SVMtrip from HN consensus sequence.

Rank	Location	Epitope (s)	Score
1	498 - 517	RLNPVSAVFDDISRSRVTRV	1
2	408 - 427	GTSHFLYQRGSSYFSPALLY	0.815
3	256 - 275	ETEEEDYKSVTPTSMVHGRL	0.767
4	99 - 118	TESIIMNAITSLSYQINGAA	0.659
5	68 - 87	DKVTSLLSSSQDVIDRIYKQ	0.659
6	284 - 303	KDLDTTALFKDWVANYPGVG	0.607
7	159 - 178	HLNFIPAPTTGSGCTRIPSF	0.588

4.2.2. Prediction of B cell Epitopes from F Protein

Table 4.4: B cell epitopes predicted by ABCpred from F consensus sequence. (Threshold 0.51)

Rank	**Sequence**	**Start position**	**Score**
1	KESIAATNEAVHEVTD	155	0.95
2	TSACMYSKTEGALTTP	367	0.93
3	LTTPYMALKGSVIANC	379	0.89
4	TGSIIVKLLPNMPKDK	58	0.88
5	GITLRLSGEFDATYLK	431	0.87
6	SPGIYSCLSGNTSACM	356	0.86
7	TLSVSTTKGFASALVP	305	0.85
7	SGLITGYPILYDSQTQ	266	0.85
7	ALYNLAGGNMDYLLTK	237	0.85
8	TTCRCADPPGIISQNY	397	0.84
9	YICLTSSLDGRPLAAA	25	0.83
9	TQLTIQALYNLAGGNM	231	0.83
9	DGLSQLSVAVGKMQQF	170	0.83

9	QGSVATSGGRRQKRFI	103	0.83
10	GSVIEELDTSYCIESD	327	0.82
10	LDCIKITQQVGVELNL	197	0.82
11	EACAKAPLEAYNRTLT	74	0.81
11	VSLIDRHSCNVLSLDG	416	0.81
12	TLTTLLTPLGDSIRKI	87	0.8
12	TGNLDISTELGNVNNS	458	0.8
13	TESNSKLDKVNVRLTS	482	0.79
13	AGIVVTGDKAVNVYTS	40	0.79
13	LGIQVNLPSVGNLNNM	283	0.79
13	SVAVGKMQQFVNDQFN	176	0.79
14	SNALDKLTESNSKLDK	475	0.77
14	TRIVTFPMSPGIYSCL	348	0.77
14	LPSVGNLNNMRATYLE	289	0.77
15	PPGIISQNYGEAVSLI	404	0.76
16	TLLWLGNNTLDQMRAT	535	0.74
17	LYLTELTTVFGPQITS	212	0.73
18	TPLGDSIRKIQGSVAT	93	0.71
18	KLLPNMPKDKEACAKA	64	0.71
18	SGGRRQKRFIGAVIGS	109	0.71
19	SGEFDATYLKNISILD	437	0.7
19	GAVIGSVALGVATAAQ	119	0.7
20	QQVGVELNLYLTELTT	204	0.68
20	MGSKPSTRIPVPLMLI	001	0.68
21	LSLGLACYLMYKQKAQ	517	0.67
21	CIESDLDLYCTRIVTF	338	0.67

21	LYDSQTQLLGIQVNLP	275	0.67
21	GGNMDYLLTKLGVGNN	243	0.67
22	AQITAAAALIQANQNA	133	0.65
23	NMRATYLETLSVSTTK	297	0.64
24	LITRIMLILSYICLTS	15	0.63
25	VYTSSQTGSIIVKLLP	52	0.62
25	FVNDQFNNTARELDCI	185	0.62
26	LTSTSALITYIALTVI	495	0.6
27	AANILRLKESIAATNE	148	0.59
28	GDKAVNVYTSSQTGSI	46	0.57
29	NISILDSQVIVTGNLD	447	0.56
30	AALIQANQNAANILRL	139	0.53

Table 4.5: B cell epitopes predicted by SVMtrip from F protein consensus sequence.

Rank	Location	Epitope	Score
1	319 - 338	VPKVVTQVGSVIEELDTSYC	1
2	384 - 403	MALKGSVIANCKITTCRCAD	0.865
3	129 - 148	VATAAQITAAAALIQANQNA	0.841
4	412 - 431	YGEAVSLIDRHSCNVLSLDG	0.812
5	50 - 69	VNVYTSSQTGSIIVKLLPNM	0.778
6	291 - 310	SVGNLNNMRATYLETLSVST	0.562
7	163 - 182	EAVHEVTDGLSQLSVAVGKM	0.508

4.3. Predicted T cell Epitopes

T cell epitopes were predicted by using HLApred and these epitopes were searched within the consensus sequences of HN and F protein. A total of 105 T cell epitopes (Table 4.6) of HN and 100 T cell epitopes of F genes (Table 4.7) were predicted. Both classes of HLA alleles were selected.

4.3.1. Prediction of T cell Epitopes from HN Protein

Table 4.6: T cell epitopes predicted by HLApred from HN consensus sequence

Sr no.	Epitopes	Length
1	AKSSYKPGR	9
2	ALESPLALL	9
3	ALGVLRTSA	9
4	ASARCPNSC	9
5	ATGKVFFST	9
6	AYQEHLNFI	9
7	DLDTTALFK	9
8	EEEDYKSVT	9
9	EGRILTVGT	9
10	EKDLDTTAL	9
11	ESIIMNAIT	9
12	ESPLALLNT	9
13	ETEEEDYKS	9
14	FHRNHTLRG	9
15	FIPAPTTGS	9
16	FKVVKTNKT	9
17	FLYQRGSSY	9
18	FNAFTRPGS	9
19	FRIAVLLLM	9
20	FRIAVLLLM	9
21	FSTLRSINL	9
22	GKELIVDDT	9
23	GKVFFSTLR	9
24	HLNFIPAPT	9
25	HQYLALGVL	9
26	IAISKTEDK	9
27	IIMNAITSL	9
28	ILAISAAAL	9
29	ILCSKVTET	9
30	ILSGCRDHS	9
31	ILSIKVSTS	9
32	IMILAISAA	9
33	IMNAITSLS	9
34	INLDDTQNR	9
35	IPAPTTGSG	9
36	IPSFDMSTT	9
37	IRMAKSSYK	9
38	ISAAALAYS	9
39	ISKTEDKVT	9
40	ISNKTATLH	9
41	IVPLLVEIL	9
42	IYKQVALES	9
43	KQVALESPL	9
44	LAISAAALA	9

45	LALLNTESI	9
46	LFKDWVANY	9
47	LGVLRTSAT	9
48	LHSPYTFNA	9
49	LIFHRNHTL	9
50	LIVDDTSDV	9
51	LKPNSPSDT	9
52	LLLMIMILA	9
53	LLMIMILAI	9
54	LMGAEGRIL	9
55	LMIMILAIS	9
56	LNFIPAPTT	9
57	LRGVFGTML	9
58	LRTSATGKV	9
59	LSIAEISNT	9
60	LTIPPNTIT	9
61	LVFRIAVLL	9
62	MILAISAAA	9
63	MIMILAISA	9
64	MLENEEREA	9
65	MLTIPPNTI	9
66	MSRAVNRVM	9
67	NEEREAKNT	9
68	NKTATLHSP	9
69	NLDDTQNRK	9
70	PLALLNTES	9
71	PLIFHRNHT	9
72	REAKNTWRL	9
73	RKSCSVSAT	9
74	RLNPVSAVF	9
75	RLVFRIAVL	9
76	RVMLENEER	9
77	SLGEDPMLT	9
78	SPALLYPMT	9
79	TKAAYTTST	9
80	TPRDLTGIS	9
81	TSTCFKVVK	9
82	TWRLVFRIA	9
83	VALESPLAL	9
84	VFRIAVLLL	9
85	VKTNKTYCL	9
86	VLLLMIMIL	9
87	VLRTSATGK	9
88	VYGGLKPNS	9
89	VYTDPYPLI	9
90	WRLVFRIAV	9
91	WVANYPGVG	9
92	YCLSIAEIS	9
93	YGGLKPNSP	9
94	YHEKDLDTT	9
95	YIGGIGKEL	9
96	YKPGRFGGK	9
97	YKQVALESP	9
98	YKSVTPTSM	9
99	YLALGVLRT	9
100	YPLIFHRNH	9
101	YPMTISNKT	9
102	YQINGAANN	9
103	YQRGSSYFS	9
104	YTHNVILSG	9

105	YVIYKRYND	9

4.3.2. Prediction of T cell Epitopes from F Protein

Table 4.7: T cell epitopes predicted by HLApred from F consensus sequence.

Sr no.	Epitopes	Length
1	AALIQANQN	9
2	AAQITAAAA	9
3	AKAPLEAYN	9
4	ALDKLTESN	9
5	ALGVATAAQ	9
6	ALIQANQNA	9
7	ALITYIALT	9
8	ALTQLTIQA	9
9	ALYNLAGGN	9
10	APLEAYNRT	9
11	AQITAAAAL	9
12	ASALVPKVV	9
13	ATSGGRRQK	9
14	DKEACAKAP	9
15	DKLTESNSK	9
16	DLYCTRIVT	9
17	DQMRATTRA	9
18	EAYNRTLTT	9
19	ELGNVNNSI	9
20	ELNLYLTEL	9
21	ETLSVSTTK	9
22	FFGVLSLGL	9
23	FGVLSLGLA	9
24	FPMSPGIYS	9
25	FVNDQFNNT	9
26	GDKAVNVYT	9
27	GEFDATYLK	9
28	GKMQQFVND	9
29	GLACYLMYK	9
30	GPQITSPAL	9
31	GSVALGVAT	9
32	GSVIEELDT	9
33	GVELNLYLT	9
34	IKITQQVGV	9
35	ILDSQVIVT	9
36	ILRLKESIA	9
37	ILYDSQTQL	9
38	IMLILSYIC	9
39	IPVPLMLIT	9
40	IRKIQGSVA	9
41	ISILDSQVI	9
42	ISLFFGVLS	9
43	ITQQVGVEL	9
44	ITSPALTQL	9

45	IVKLLPNMP	9	71	MLILSYICL	9
46	IVTGNLDIS	9	72	MLITRIMLI	9
47	KDKEACAKA	9	73	MPKDKEACA	9
48	KEACAKAPL	9	74	MYKQKAQQK	9
49	KLTESNSKL	9	75	NLYLTELTT	9
50	LAAAGIVVT	9	76	PKDKEACAK	9
51	LDGRPLAAA	9	77	PLEAYNRTL	9
52	LGIQVNLPS	9	78	QKAQQKTLL	9
53	LGVATAAQI	9	79	QKTLLWLGN	9
54	LIGSGLITG	9	80	QNYGEAVSL	9
55	LILSYICLT	9	81	QQKTLLWLG	9
56	LKGSVIANC	9	82	RHSCNVLSL	9
57	LKNISILDS	9	83	RKIQGSVAT	9
58	LLPNMPKDK	9	84	RLKESIAAT	9
59	LLWLGNNTL	9	85	RLSGEFDAT	9
60	LMLITRIML	9	86	SKTEGALTT	9
61	LMYKQKAQQ	9	87	SLIGSGLIT	9
62	LNNMRATYL	9	88	TKGFASALV	9
63	LRLKESIAA	9	89	TKLGVGNNQ	9
64	LRLSGEFDA	9	90	TLDQMRATT	9
65	LSLDGITLR	9	91	TLLWLGNNT	9
66	LSLGLACYL	9	92	TLTTLLTPL	9
67	LSSLIGSGL	9	93	TNEAVHEVT	9
68	LTSTSALIT	9	94	TSACMYSKT	9
69	LTTPYMALK	9	95	TTKGFASAL	9
70	LYNLAGGNM	9	96	TYLETLSVS	9

97	VGNLNNMRA	9	99	VIGSVALGV	9
98	VGNNQLSSL	9	100	YTSSQTGSI	9

4.4. Overlapping Epitopes

Predicted T cell epitopes were searched within the predicted B cell epitopes. A total of 45 B cell epitopes (Table 4.8) of HN and 49 B (Table 4.9) cell epitopes of F genes were shortlisted containing the T cell epitopes within them. These were the selected overlapping epitopes.

3.4.1 Overlapping B and T cell Epitopes of HN Protein

Table 4.8: Overlapping B and T cell epitopes of HN protein.

Sr no.	Overlapping Epitopes
1	NEEREAKNTWRLVFRI
2	PNTITLMGAEGRILTV
3	KDLDTTALFKDWVANYPGVG
4	TGVYTDPYPLIFHRNH
5	HLNFIPAPTTGSGCTRIPSF
6	PSAYQEHLNFIPAPTT
7	DERVWFPVYGGLKPNS
8	EFRIVPLLVEILKDNR
9	PYTFNAFTRPGSVPCQ
10	TATLHSPYTFNAFTRP
11	EGRILTVGTSHFLYQR
12	VNRVMLENEEREAKNT
13	WVANYPGVGGGSFVDE
14	MSTTHYCYTHNVILSG
15	GKVFFSTLRSINLDDT
16	GEDPMLTIPPNTITLM
17	KELIVDDTSDVTSFYP
18	VFRIAVLLLMIMILAI
19	RLNPVSAVFDDISRSRVTRV
20	QDVIDRIYKQVALESP
21	YSMGTSTPRDLTGISI
22	YQINGAANNSGCGAPV
23	LDDGQARLNPVSAVFD
24	QASARCPNSCITGVYT
25	LGVLRTSATGKVFFST
26	TESIIMNAITSLSYQINGAA
27	HSHSHQYLALGVLRTS
28	VKTNKTYCLSIAEISN
29	TPLGCDILCSKVTETE
30	TGSGCTRIPSFDMSTT
31	DTAQEGKYVIYKRYND

32	IYKQVALESPLALLNT
33	GISIAISKTEDKVTSL
34	ETEEEDYKSVTPTSMVHGRL
35	AITSLSYQINGAANNS
36	KAAYTTSTCFKVVKTN
37	SYFSPALLYPMTISNK
38	LRSINLDDTQNRKSCS
39	GTSHFLYQRGSSYFSPALLY
40	AKSSYKPGRFGGKRVQ
41	GKRVQQAILSIKVSTS
42	SIKVSTSLGEDPMLTI
43	GLKPNSPSDTAQEG
44	RMAKSSYKPGRFGGKR
45	PDEQDYQIRMAKSSYK

3.4.2 Overlapping B and T cell Epitopes of F Protein

Table 4.9: Overlapping B and T cell epitopes of F protein

Sr no.	Overlapping Epitopes
1	KESIAATNEAVHEVTD
2	GAVIGSVALGVATAAQ
3	QQVGVELNLYLTELTT
4	GSVIEELDTSYCIESD
5	VATAAQITAAAALIQANQNA
6	LTSTSALITYIALTVI
7	GITLRLSGEFDATYLK
8	FVNDQFNNTARELDCI
9	AQITAAAALIQANQNA
10	TQLTIQALYNLAGGNM
11	AALIQANQNAANILRL
12	AGIVVTGDKAVNVYTS
13	AANILRLKESIAATNE
14	VPKVVTQVGSVIEELDTSYC
15	CIESDLDLYCTRIVTF
16	SGEFDATYLKNISILD
17	NISILDSQVIVTGNLD
18	TLLWLGNNTLDQMRAT
19	LITRIMLILSYICLTS
20	PPGIISQNYGEAVSLI
21	LDCIKITQQVGVELNL
22	EACAKAPLEAYNRTLT
23	ITSPALTQLTIQALYNLAGG
24	ITSPALTQLTIQALYNLAGG
25	NMRATYLETLSVSTTK
26	TLTTLLTPLGDSIRKI
27	LPSVGNLNNMRATYLE
28	ALYNLAGGNMDYLLTK
29	VSLIDRHSCNVLSLDG
30	TRIVTFPMSPGIYSCL
31	QGSVATSGGRRQKRFI
32	YGEAVSLIDRHSCNVLSLDG

33	SVGNLNNMRATYLETLSVST
34	LGIQVNLPSVGNLNNM
35	YICLTSSLDGRPLAAA
36	TPLGDSIRKIQGSVAT
37	MALKGSVIANCKITTCRCAD
38	TLSVSTTKGFASALVP
39	LTTPYMALKGSVIANC
40	FGVLSLGLACYLMYKQKAQQ

41	TSACMYSKTEGALTTP
42	TGSIIVKLLPNMPKDK
43	SVAVGKMQQFVNDQFN
44	GDKAVNVYTSSQTGSI
45	NMPKDKEACAKA
46	VYTSSQTGSIIVKLLP
47	SNALDKLTESNSKLDK
48	KLLPNMPKDKEACAKA

4.5. Evaluation of Epitopes

The epitopes were evaluated on the basis of their immunogenic and antigenic potential, their nature as allergen and their property of inducing IFN gamma response.

4.5.1. Immunogenic Potential

Using IEDB server, only 23 out of 47 shortlisted epitopes of HN protein, showed positive immunogenic value and rest were discarded. In the same way, only 22 out of 49 epitopes of F protein showed positive value for immunogenicity.

Table 4.10: Epitopes of HN protein with positive immunogenicity value.

Sr no.	Epitopes	Immunogenicity (IEDB)
1	NEEREAKNTWRLVFRI	0.46488
2	PNTITLMGAEGRILTV	0.39312
3	KDLDTTALFKDWVANYPGVG	0.3736
4	TGVYTDPYPLIFHRNH	0.34703
5	HLNFIPAPTTGSGCTRIPSF	0.29726

6	PSAYQEHLNFIPAPTT	0.29103
7	DERVWFPVYGGLKPNS	0.2529
8	EFRIVPLLVEILKDNR	0.22696
9	PYTFNAFTRPGSVPCQ	0.2235
10	TATLHSPYTFNAFTRP	0.21424
11	EGRILTVGTSHFLYQR	0.19188
12	VNRVMLENEEREAKNT	0.16169
13	WVANYPGVGGGSFVDE	0.16089
14	MSTTHYCYTHNVILSG	0.14592
15	GKVFFSTLRSINLDDT	0.14494
16	GEDPMLTIPPNTITLM	0.1431
17	KELIVDDTSDVTSFYP	0.13093
18	VFRIAVLLLMIMILAI	0.123
19	RLNPVSAVFDDISRSRVTRV	0.1114
20	QDVIDRIYKQVALESP	0.07822
21	YSMGTSTPRDLTGISI	0.02736
22	YQINGAANNSGCGAPV	0.01113
23	LDDGQARLNPVSAVFD	0.00827

Table 4.11: Epitopes of F protein with positive immunogenicity value.

Sr no.	Epitopes	Immunogenicity (IEDB)
1	KESIAATNEAVHEVTD	0.54699
2	GAVIGSVALGVATAAQ	0.29044
3	QQVGVELNLYLTELTT	0.28827
4	GSVIEELDTSYCIESD	0.28684
5	VATAAQITAAAALIQANQNA	0.23354
6	LTSTSALITYIALTVI	0.2295
7	GITLRLSGEFDATYLK	0.21108
8	FVNDQFNNTARELDCI	0.2094
9	AQITAAAALIQANQNA	0.17113
10	TQLTIQALYNLAGGNM	0.15368
11	AALIQANQNAANILRL	0.14628
12	AGIVVTGDKAVNVYTS	0.13776
13	AANILRLKESIAATNE	0.12053
14	VPKVVTQVGSVIEELDTSYC	0.1192
15	CIESDLDLYCTRIVTF	0.10253
16	SGEFDATYLKNISILD	0.09681
17	NISILDSQVIVTGNLD	0.08488
18	TLLWLGNNTLDQMRAT	0.06593

19	LITRIMLILSYICLTS	0.05875
20	PPGIISQNYGEAVSLI	0.05208
21	LDCIKITQQVGVELNL	0.04273
22	EACAKAPLEAYNRTLT	0.02907

4.5.2. Antigenic Potential

Among the shortlisted immunogenic epitopes of HN protein, 13 antigenic epitopes were identified using Vaxigen v2. In the same way, 19 antigenic epitopes were identified using Vaxigen v2 in F protein.

Table 4.12: Epitopes of HN protein showing antigenic potential.

Sr no.	**Epitopes**	**Antigenecity (Vaxigen)**
1	NEEREAKNTWRLVFRI	0.4395
2	TGVYTDPYPLIFHRNH	0.7811
3	HLNFIPAPTTGSGCTRIPSF	0.6082
4	PSAYQEHLNFIPAPTT	0.8876
5	DERVWFPVYGGLKPNS	0.5259
6	EFRIVPLLVEILKDNR	1.0018
7	EGRILTVGTSHFLYQR	0.491
8	VNRVMLENEEREAKNT	0.4317

9	MSTTHYCYTHNVILSG	0.5395
10	GKVFFSTLRSINLDDT	0.6695
11	RLNPVSAVFDDISRSRVTRV	0.4336
12	YSMGTSTPRDLTGISI	0.02736
13	LDDGQARLNPVSAVFD	0.00827

Table 4.13: Epitopes of F protein showing antigenic potential.

Sr no.	**Epitopes**	**Antigenecity (Vaxigen)**
1	KESIAATNEAVHEVTD	0.5019
2	GAVIGSVALGVATAAQ	0.821
3	QQVGVELNLYLTELTT	0.8057
4	GSVIEELDTSYCIESD	0.4043
5	VATAAQITAAAALIQANQNA	0.5345
6	LTSTSALITYIALTVI	0.7594
7	GITLRLSGEFDATYLK	0.8478
8	FVNDQFNNTARELDCI	0.8196
9	AQITAAAALIQANQNA	0.5261
10	TQLTIQALYNLAGGNM	0.7237
11	AGIVVTGDKAVNVYTS	0.4117

12	AANILRLKESIAATNE	0.5636
13	VPKVVTQVGSVIEELDTSYC	0.3474
14	CIESDLDLYCTRIVTF	0.8964
15	SGEFDATYLKNISILD	1.0832
16	NISILDSQVIVTGNLD	0.8809
17	LITRIMLILSYICLTS	0.6397
18	PPGIISQNYGEAVSLI	0.7655
19	LDCIKITQQVGVELNL	1.2511

4.5.3. Allergenecity Prediction

Using AllergenFP, 7 HN epitopes were further identified as non-allergen and rest were discarded. 14 F epitopes were identified as non-allergen. Rest of them were discarded.

Table 4.14: HN protein epitopes with non-allergen nature.

Sr no.	**Epitopes**	**Allergenecity (Allergen FP)**
1	HLNFIPAPTTGSGCTRIPSF	Non- Allergen
2	EGRILTVGTSHFLYQR	Non-Allergen
3	VNRVMLENEEREAKNT	Non Allergen
4.	MSTTHYCYTHNVILSG	Non-Allergen

5	GKVFFSTLRSINLDDT	Non-Allergen
6	RLNPVSAVFDDISRSRVTRV	Non-Allergen
7	LDDGQARLNPVSAVFD	Non-Allergen

Table 4.15: F protein epitopes with non-allergen nature.

Sr no.	**Epitopes**	**Allergenecity (Allergen FP)**
1	KESIAATNEAVHEVTD	Non-Allergen
2	GAVIGSVALGVATAAQ	Non-Allergen
3	QQVGVELNLYLTELTT	Non-Allergen
4	GSVIEELDTSYCIESD	Non-Allergen
5	VATAAQITAAAALIQANQNA	Non-Allergen
6	LTSTSALITYIALTVI	Non-Allergen
7	GITLRLSGEFDATYLK	Non-Allergen
8	FVNDQFNNTARELDCI	Non-Allergen
9	AGIVVTGDKAVNVYTS	Non-Allergen
10	AANILRLKESIAATNE	Non-Allergen
11	VPKVVTQVGSVIEELDTSYC	Non-Allergen
12	CIESDLDLYCTRIVTF	Non-Allergen

13	PPGIISQNYGEAVSLI	Non-Allergen
14	LDCIKITQQVGVELNL	Non-Allergen

4.5.4. Conservance

Blastp was used to find the conservance of the epitopes and all epitopes of both HN protein and F protein were found to be 100 % conserved as shown in table 4.16 and 4.17.

HN Protein Epitopes:

Table 4.16: Conservancy of HN protein epitopes.

Sr no.	**Epitopes**	**Conservance (Blastp)**
1	EGRILTVGTSHFLYQR	100%
2	GKVFFSTLRSINLDDT	100%
3	RLNPVSAVFDDISRSRVTRV	100%

F Protein Epitopes:

Table 4.17: Conservancy of F protein epitopes.

Sr no.	**Epitopes**	**Conservance (Blastp)**
1	KESIAATNEAVHEVTD	100%
2	QQVGVELNLYLTELTT	100%
3	GSVIEELDTSYCIESD	100%
4	VATAAQITAAAALIQANQNA	100%
5	PPGIISQNYGEAVSLI	100%

4.6. Predicted IFN-Gamma Inducing Epitopes

Using IFNepitope, only 3 epitopes of HN protein and 5 epitopes of F protein were IFN-gamma inducing epitopes. Rest of the epitopes were discarded.

HN Protein Epitopes:

Table 4.18: IFN gamma inducing HN protein epitopes.

Sr no.	Epitopes	IFNepitope
1	EGRILTVGTSHFLYQR	Positive
2	GKVFFSTLRSINLDDT	Positive
3	RLNPVSAVFDDISRSRVTRV	Positive

F Protein Epitopes:

Table 4.19: IFN gamma inducing F protein epitopes.

Sr no.	Epitopes	IFNepitope
1	KESIAATNEAVHEVTD	Positive
2	QQVGVELNLYLTELTT	Positive
3	GSVIEELDTSYCIESD	Positive
4	VATAAQITAAAALIQANQNA	Positive
5	PPGIISQNYGEAVSLI	Positive

4.7. Finalized Epitopes

Hence, the 3 finalized epitopes of HN protein (Table 4.19) and 5 finalized epitopes of F protein (Table 4.20) showed positive immunogenic value, high antigenic threshold, and positive IFN gamma response and were non-allergen in nature.

HN Protein Epitopes:

Table 4.19: Properties of finalized epitopes in HN Protein.

Sr no	Epitopes	Immunogenicity	Antigenecity	Allergenicity	IFNepitope	Conservancy
1	EGRILTVGTSHFLYQR	0.19188	0.491	Non Allergen	Positive	100%
2	GKVFFSTLRSINLDDT	0.14494	0.6695	Non Allergen	Positive	100%
3	RLNPVSAVFDDISRSRVTRV	0.1114	0.4336	Non Allergen	Positive	100%

F Protein Epitopes:

Table 4.20: Properties of finalized epitopes in F Protein.

Sr no	Epitopes	Immunogenicity	Antigenecity	Allergenicity	IFNepitope	Conservancy
1	KESIAATNEAVHEVTD	0.54699	0.5019	Non Allergen	Positive	100%
2	QQVGVELNLYLTELTT	0.28827	0.8057	Non Allergen	Positive	100%
3	GSVIEELDTSYCIESD	0.28684	0.4043	Non Allergen	Positive	100%
4	VATAAQITAAAALIQANQNA	0.23354	0.5345	Non Allergen	Positive	100%
5	PPGIISQNYGEAVSLI	0.05208	0.7655	Non Allergen	Positive	100%

4.8. Multi-Epitope Vaccine Design

4.8.1. Epitope-Epitope Interactions and Arrangement

The finalized epitopes mentioned in table 8 and 9 were analyzed on the basis of their binding affinity with each other. Using Haddock server, the initial pair of epitopes showing maximum binding energy were selected and joined with flexible linkers GPGPG. The arrangement (HN2-F1) was considered the best arrangement based on the refinement score of HADDOCK. Now, after joining them with linker, these were further analyzed for their binding affinity with the rest of 6 epitopes. (HN2-F1-HN3) showed the best refinement score. (HN2-F1-HN3-HN1-F3-F2-F5-F4) was the finalized construct obtained through epitope-epitope interactions. All the possible combinations along with their refinement score are shown in table 4.21.

Table 4.21: Epitope-Epitope interactions and their arrangements of initial epitopes for vaccine design.

	EGRILTVGTSHFLYQR	GKVFFSTLRSINLDDT	RLNPVSAVFDDISRSRVTRV	KESIAATNEAVHEVTD	QQVGVELNLYLTELTT	GSVIEELDTSYCIESD	VATAAQITAAAALIQANQNA	PPGIISQNYGEAVSLI
EGRILTVGTSHFLYQR		-85.9 +/- 1.7	-79.2 +/- 0.2	-97.8 +/- 0.5	-80.2 +/- 0.8	-68.5 +/- 2.4	-56.7 +/- 0.9	-81.0 +/- 1.5
GKVFFSTLRSINLDDT	-77.3 +/- 0.4		-87.0 +/- 1.9	-99.6 +/- 2.0	-83.0 +/- 1.6	-85.6 +/- 1.3	-56.8 +/- 0.9	-75.8 +/- 1.5
RLNPVSAVFDDISRSRVTRV	-80.8 +/- 0.5	-75.9 +/- 1.5		-84.9 +/- 1.6	-82.0 +/- 1.0	-79.3 +/- 1.8	-69.8 +/- 1.6	-84.8 +/- 0.4
KESIAATNEAVHEVTD	-94.2 +/- 0.6	-87.6 +/- 1.0	-71.9 +/- 0.4		-70.9 +/- 1.8	-61.1 +/- 1.1	-51.7 +/- 1.1	-63.4 +/- 1.4
QQVGVELNLYLTELTT	-66.8 +/- 0.3	-73.2 +/- 1.1	-78.8 +/- 1.8	-60.7 +/- 1.0		-60.2 +/- 2.2	-50.3 +/- 1.3	-66.1 +/- 1.2
GSVIEELDTSYCIESD	-88.2 +/- 0.5	-70.1 +/- 1.0	-76.3 +/- 0.7	-67.1 +/- 0.7	-58.1 +/- 1.1		-51.7 +/- 0.9	-71.6 +/- 1.7
VATAAQITAAAALIQANQNA	-63.7 +/- 0.4	-52.2 +/- 1.0	-71.3 +/- 1.1	-56.0 +/- 2.2	-66.5 +/- 0.6	-45.8 +/- 0.9		-66.5 +/- 0.1
PPGIISQNYGEAVSLI	-83.1 +/- 0.3	-86.8 +/- 1.1	-92.2 +/- 1.1	-62.6 +/- 1.0	-58.6 +/- 1.1	-62.4 +/- 0.3	-63.6 +/- 0.5	

As shown in table 4.21, the epitopes **GKVFFSTLRSINLDDT** and **KESIAATNEAVHEVTD** showed maximum HADDOCK refinement score i.e. **-99.6+/-2.0** . So, these two epitopes were selected to be the first two epitopes in the multiepitope arrangement of vaccine construct. The complex of these two epitopes will be docked against all the remaining epitopes and the epitopes are arranged accordingly. The results are shown in the table 4.22.

Table 4.22: Epitope-Epitope interactions and their arrangements for vaccine design.

Best Epitope Combination	Other Epitopes	HADDOCK REFINEMENT SCORE
Complex 1		
E1xE2	EGRILTVGTSHFLYQR	-86.3 +/- 1.6
E1xE2	RLNPVSAVFDDISRSRVTRV	-93.2 +/- 1.2
E1xE2	QQVGVELNLYLTELTT	-61.0 +/- 1.0
E1xE2	GSVIEELDTSYCIESD	-80.1 +/- 1.5
E1xE2	VATAAQITAAAALIQANQNA	-58.6 +/- 0.5
E1xE2	PPGIISQNYGEAVSLI	-87.8 +/- 0.5
Complex 2		
E1xE2xE3	EGRILTVGTSHFLYQR	-88.2 +/- 0.8
E1xE2xE3	QQVGVELNLYLTELTT	-83.5 +/- 2.4
E1xE2xE3	GSVIEELDTSYCIESD	-67.6 +/- 1.4
E1xE2xE3	VATAAQITAAAALIQANQNA	-69.1 +/- 1.6
E1xE2xE3	PPGIISQNYGEAVSLI	-68.8 +/- 1.4
Complex 3		
E1xE2xE3xE4	QQVGVELNLYLTELTT	-85.2 +/- 2.6
E1xE2xE3xE4	GSVIEELDTSYCIESD	-86.7 +/- 2.8
E1xE2xE3xE4	VATAAQITAAAALIQANQNA	-83.7 +/- 3.3
E1xE2xE3xE4	PPGIISQNYGEAVSLI	-70.2 +/- 3.2
Complex 4		
E1xE2xE3xE4xE5	QQVGVELNLYLTELTT	-67.7 +/- 2.1
E1xE2xE3xE4xE5	VATAAQITAAAALIQANQNA	-53.6 +/- 1.7
E1xE2xE3xE4xE5	PPGIISQNYGEAVSLI	-66.9 +/- 1.6
Complex 5		
E1xE2xE3xE4xE5xE6	VATAAQITAAAALIQANQNA	-59.8 +/- 1.5
E1xE2xE3xE4xE5xE6	PPGIISQNYGEAVSLI	-113.5 +/- 0.8
Complex 6		
E1xE2xE3xE4xE5xE6xE7	VATAAQITAAAALIQANQNA	-62.9 +/- 2.1

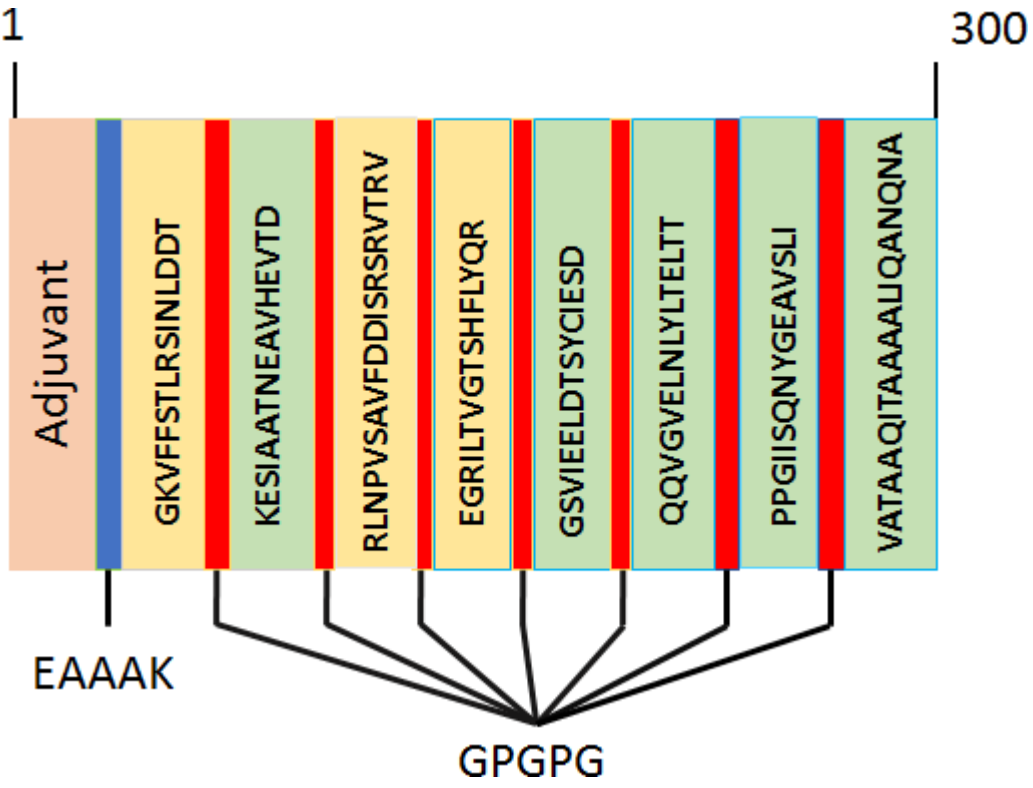
4.8.2. Addition of Adjuvant

The finalized arrangement of epitopes was linked with flexible GPGPG linkers. Furthermore, an adjuvant (Accession id: WP_000593522) Cholera toxin b (CTB) of 124 amino-acid length;

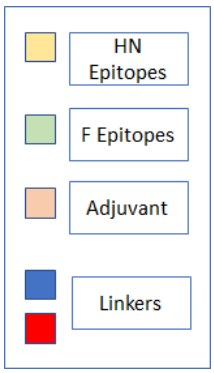
(MIKLKFGVFFTVLLSSAYAHGTPQNITDLCAEYHNTQIYTLNDKIFSYTESLAGK REMAIITFKNGAIFQVEVPGSQHIDSQKKAIERMKDTLRIAYLTEAKVEKLCVWN NKTPHAIAAISMAN) was attached at the N terminal of the epitopes via EAAAK linker.

The final length of the vaccine construct, including adjuvant and linkers, was 300 amino acids.

(a)



(c)



(b) **MIKLKFGVFFTVLLSSAYAHGTPQNITDLCAEYHNTQIYTLNDKIFS YTESLAGKREMAIITFKNGAIFQVEVPGSQHIDSQKKAIERMKDTLRIA YLTEAKVEKLCVWNNKTPHAIAAISMANEAAAKGKVFFSTLRSINLD DTGPGPGKESIAATNEAVHEVTDGPGPGRLNPVSAVFDDISRSRVTR VGPGPGEGRILTVGTSHFLYRQGPGPGGSVIEELDTSYCIESDGPGPG QQVGVELNLYLTELTTGPGPGPPGIISQNYGEAVSLIGPGPGVATAAQI TA AAALIQANQNA**

Figure 4.2 (a) Schematic diagram of proposed vaccine construct. **(b)** Overall sequence of the construct. **(c)** Color key for the vaccine construct.

4.9. Physiochemical Properties of Vaccine Construct

Molecular weight of the finalized vaccine construct was estimated to be 31.68285 kD. Proteins with molecular weight smaller than 110 kD are considered to be potent vaccine candidates. The instability index (II) of the construct was computed to be 27.72 (>40), classified it as stable protein. The aliphatic index was computed to be 72.09, indicating that the construct is thermostable. The grand average of hydropathicity (GRAVY) score was -0.134, pointing out that the protein is hydrophilic in nature and can have good interaction with surrounding water molecules.

4.10. Modeling of the 3D Structure of Vaccine Construct

ITasser online server was used to model the 3D structure of the construct. The structure was modeled on the basis of its homology with several templates. The predicted model with the highest C score (-3.16) was selected (Figure 4.3).

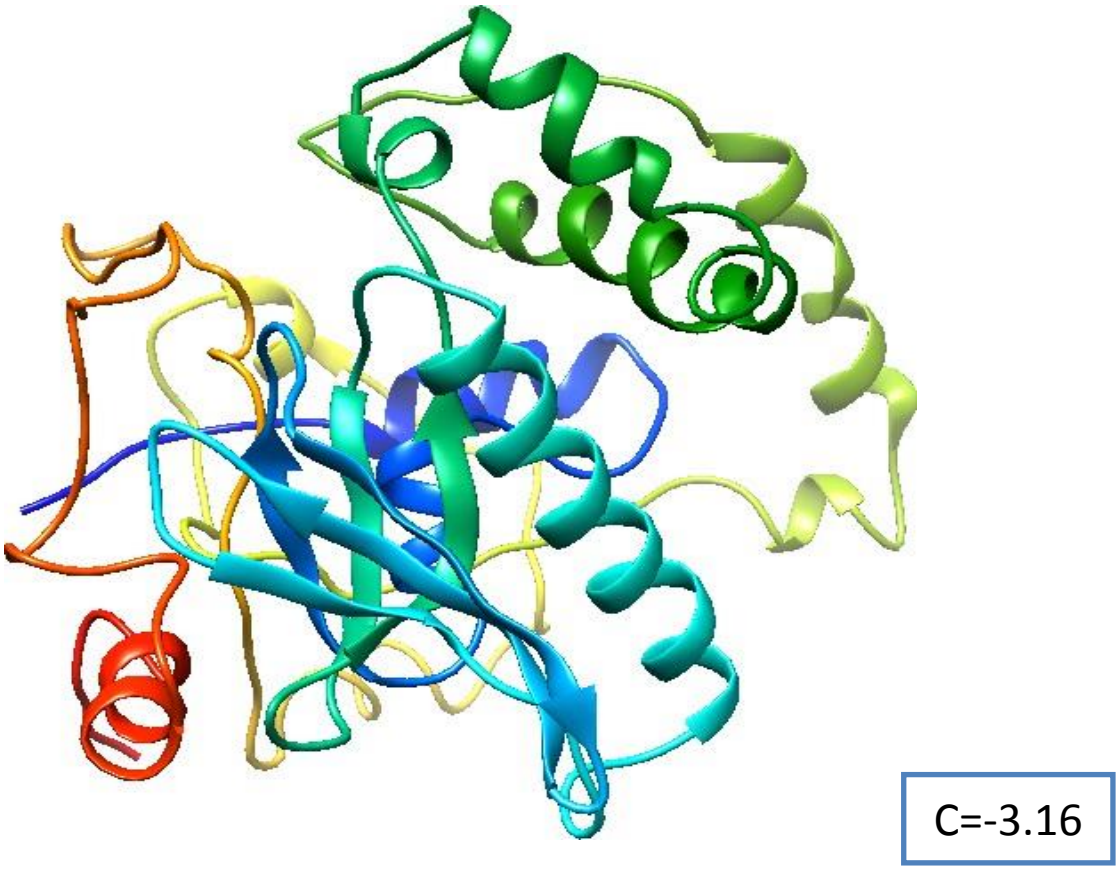
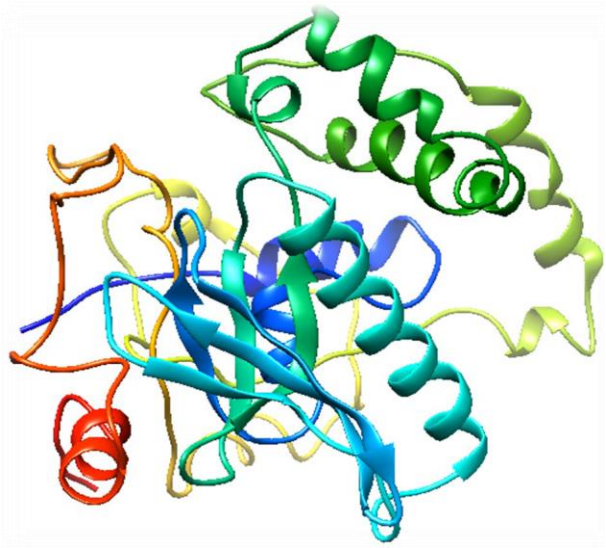


Figure 4.3: 3D structure of the vaccine construct made by ITasser.

4.11. Refinement and Validation of 3D Modeled Vaccine Construct

Galaxy Refine online tool was used to process the predicted model into a more refined version. The best model was preceded further and was validated through various online tools.

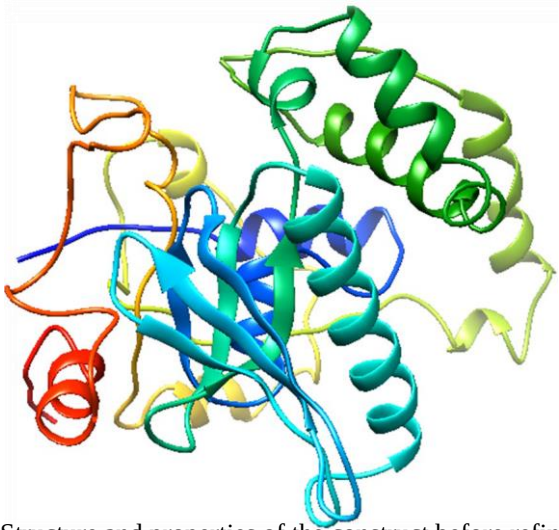
(a)



RMSD 0
Poor Rotamers **1.4**
Rama favored **73.5**

(Galaxy Refine)

(b)



RMSD 9742
Poor Rotamers **0.0**
Rama favored **92.6**

Figure 4.4: (a) Structure and properties of the construct before refining **(b)** Structure and properties after refining.

4.12. Validation of 3D structure

The refined structure was validated through various online tools. Ramachandran analysis through RAMPAGE server showed, 92.6% of the structure was under favorable region, 9.4% was under the allowed region, and 2.0% was observed under the disallowed region regarding it as a high quality structure (Figure 4.5). ProSA-web showed a Z-score of −6.5, which lies inside the range of acceptable scores. Online tool Verify 3D also evaluated the 3D structure as pass, signaling out its good quality.

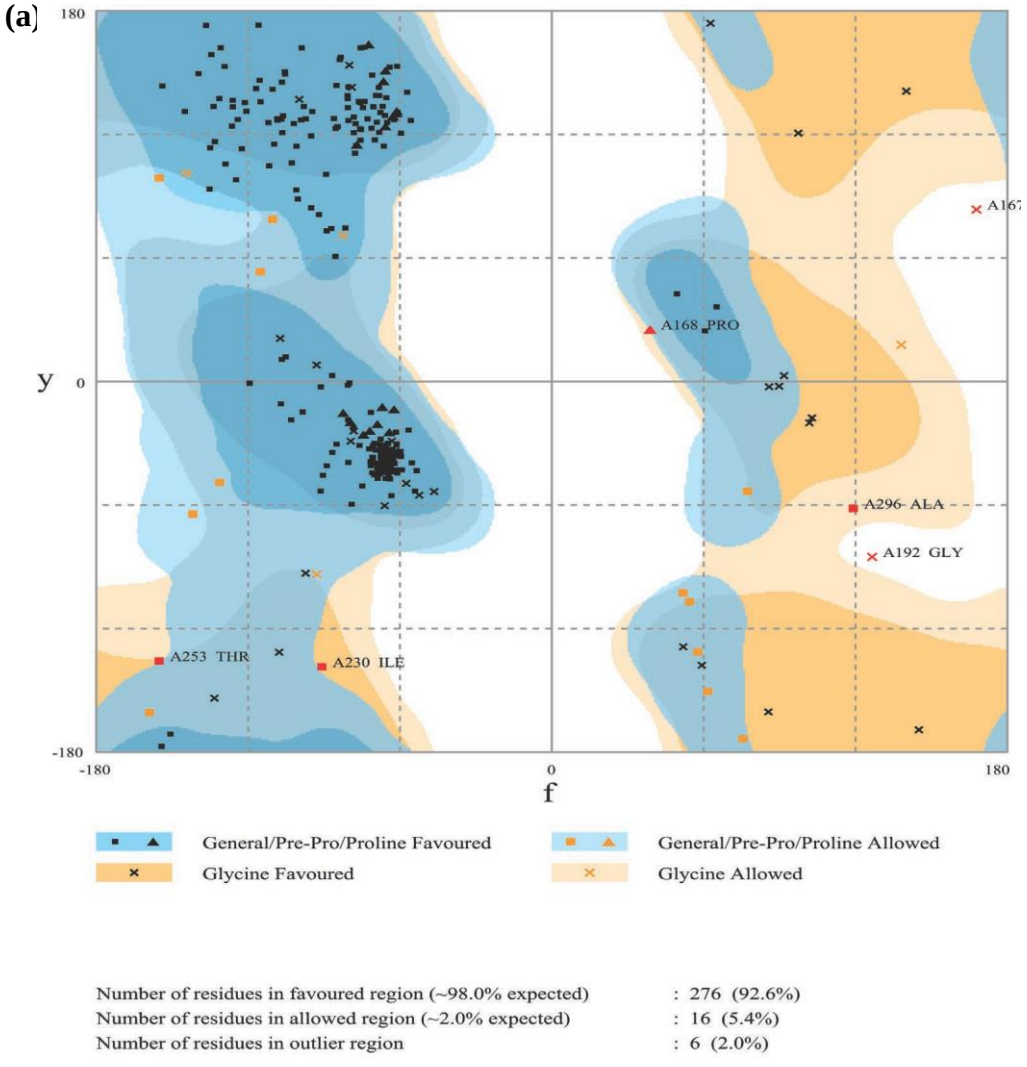


Figure 4.5: Ramachandran plot analysis.

(b)

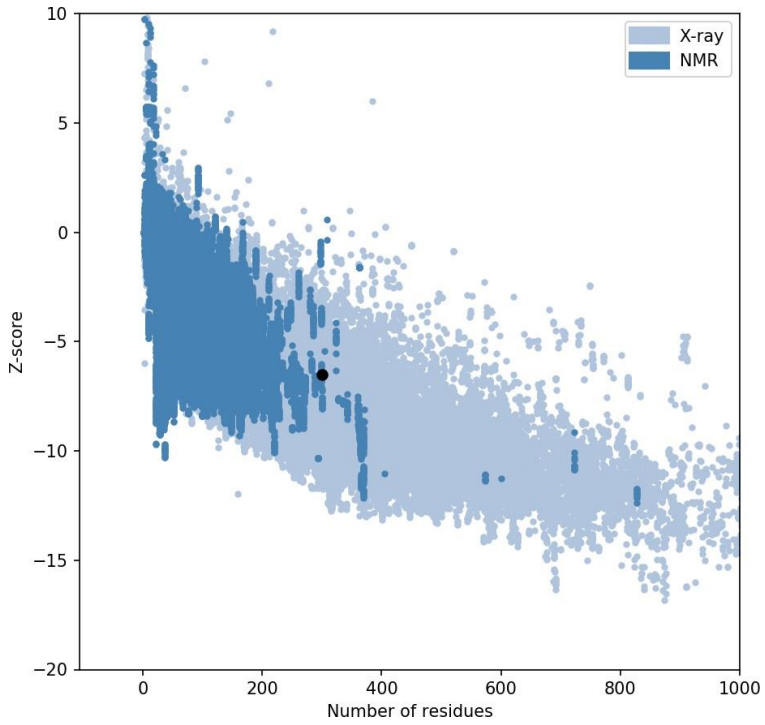


Figure 4.6: ProSA-web-Z score analysis.

(c)

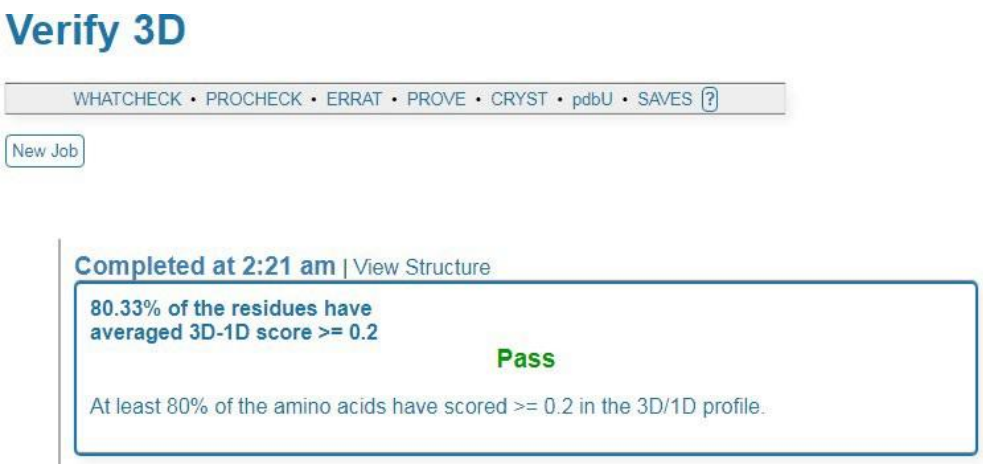
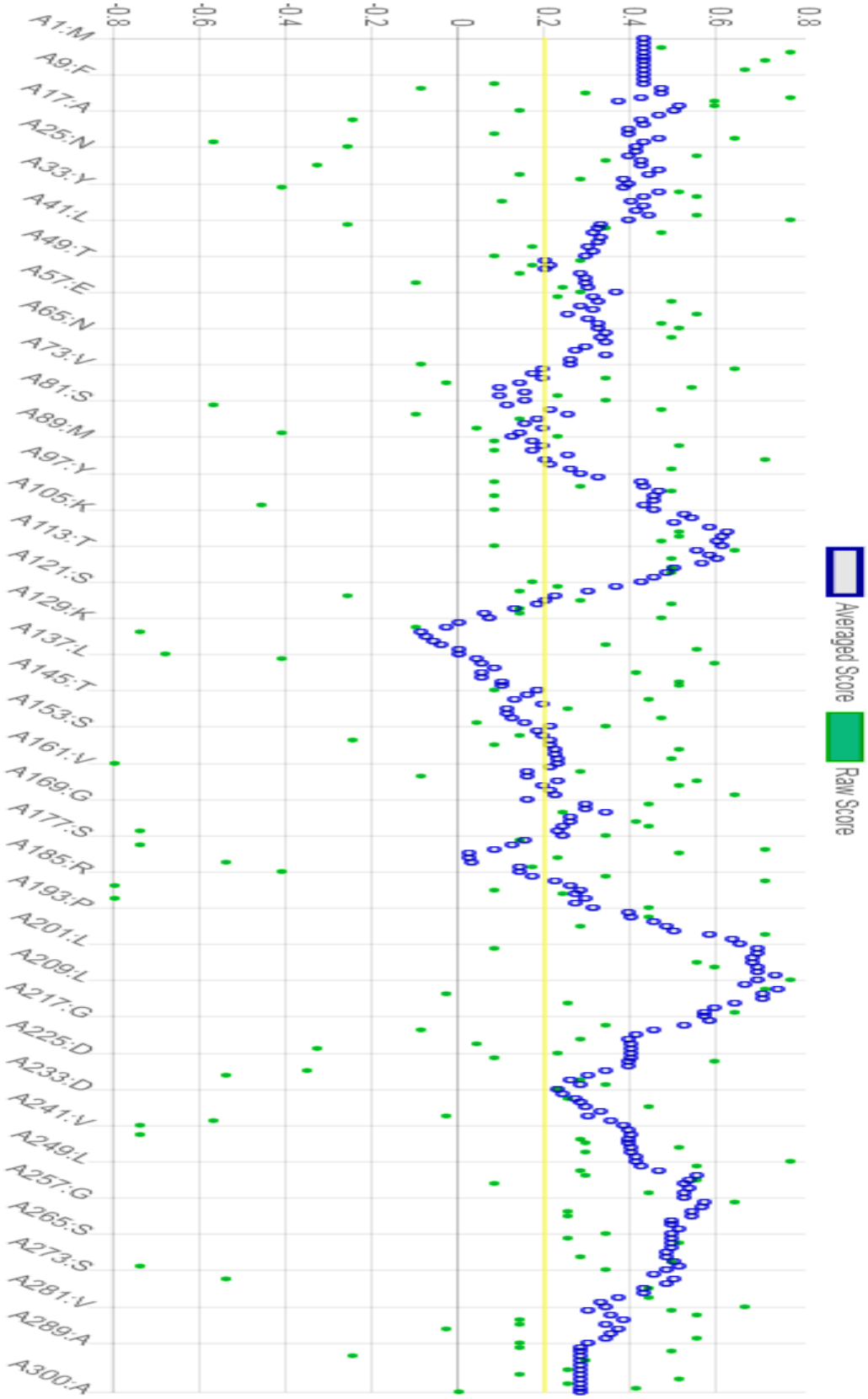


Figure 4.7: Verify 3D verification of 3d stucture.


-0.8
-0.6
-0.4
-0.2
0
0.2
0.4
0.6
0.8
A1:M
A9:F
A17:A
A25:N
A33:Y
A41:L
A49:T
A57:E
A65:N
A73:V
A81:S
A89:M
A97:Y
A105:K
A113:T
A121:S
A129:K
A137:L
A145:T
A153:S
A161:V
A169:G
A177:S
A185:R
A193:P
A201:L
A209:L
A217:G
A225:D
A233:D
A241:V
A249:L
A257:G
A265:S
A273:S
A281:V
A289:A
A300:A
Averaged Score
Raw Score

4.13. Molecular Docking of the vaccine with Toll like Receptors

The docking analysis of the vaccine construct and toll like receptors was carried out by HADDOCK server. The construct was docked separately with three toll like receptors (TLR2-type1, TLR2-type2, and TLR4) of chicken (*Gallus gallus*). The docking analysis showed a great binding affinity of the vaccine construct with all three of the receptors. The docking output is showed as clusters and the top cluster with the minimum Z-score is considered to be most reliable. Z-score indicates how many standard deviations from the average this cluster is located in terms of score.

In case of TLR 4, HADDOCK clustered 398 structures in 3 clusters, which represent 99.5 % of the water-refined models HADDOCK generated. Top cluster was selected with the Z score of -1.5 (Fig 4.8a). Further statistical data are given in the table. PDBsum online server showed detailed interaction analysis; 48 interface residues of the construct associated with 55 residues of TLR4 (Fig 4.10a).The interface area (Å2) of the vaccine was 2487, while that of TLR4 was found to be 2324. A total of 4 salt bridges, 23 hydrogen bonds and 357 non bonded contacts were found to be formed between vaccine and TLR4.

In case of TLR2 type 1, HADDOCK clustered 204 structures in 22 clusters, which represent 51.0 % of the water-refined models HADDOCK generated. Top cluster was selected with the Z score of -1.9 (Fig 4.8b). Further statistical data are given in the table. PDBsum online server showed detailed interaction analysis; 41 interface residues of the construct associated with 54 residues of TLR4 (Fig 4.10b).The interface area (Å2) of the vaccine was 2653, while that of TLR4 was found to be 2390. A total of 1 salt bridge, 21 hydrogen bonds and 324 non bonded contacts were found to be formed between vaccine and TLR2 type1.

In case of TLR 2 type 2, HADDOCK clustered 255 structures in 32 clusters, which represent 63.75 % of the water-refined models HADDOCK generated. Top cluster was selected with the Z score of -1.4 (Fig 4.8c) Further statistical data are given in the table. PDBsum online server showed detailed interaction analysis; 51 interface residues of the construct associated with 60 residues of TLR4 (Fig 4.10c).The interface area (Å2) of the vaccine was 2788, while that of TLR4 was found to be 2651. A total of 2 salt bridges, 21

hydrogen bonds and 331 non bonded contacts were found to be formed between vaccine and TLR2 type2.

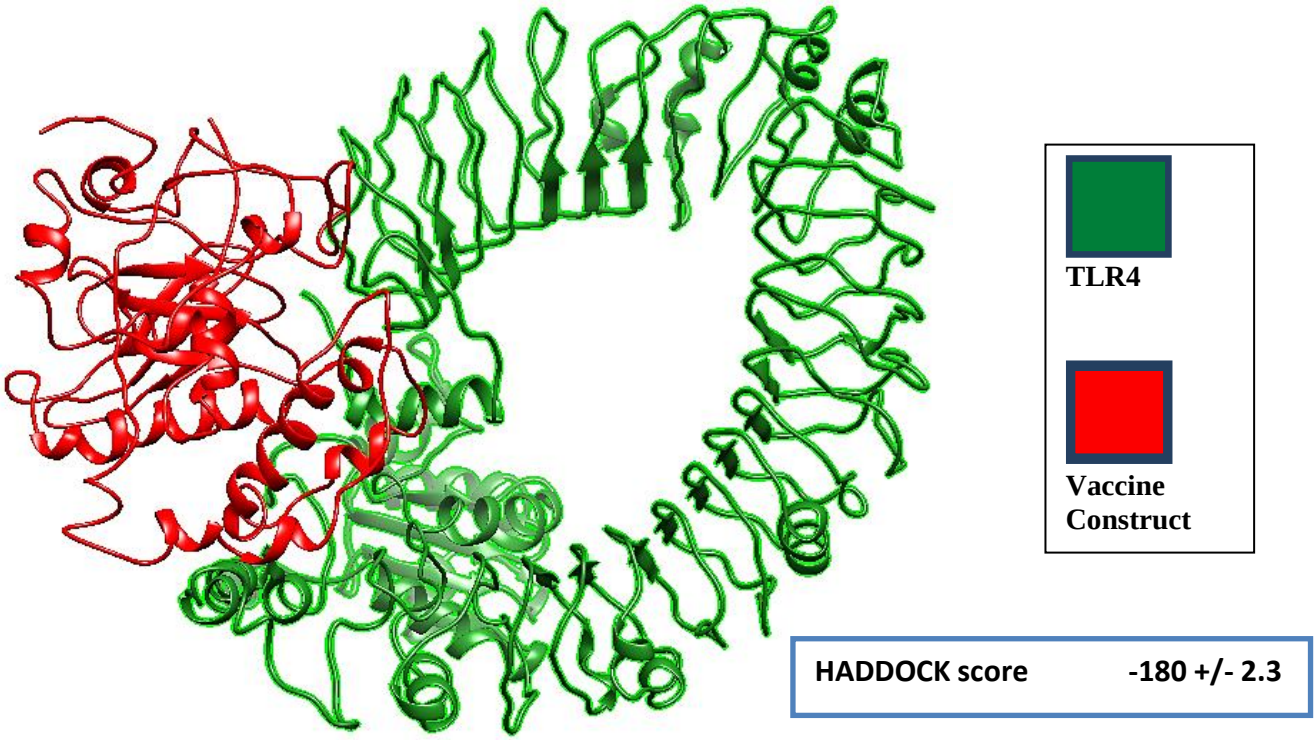


Figure 4.8 (a): Interaction of Vaccine construct with TLR4.

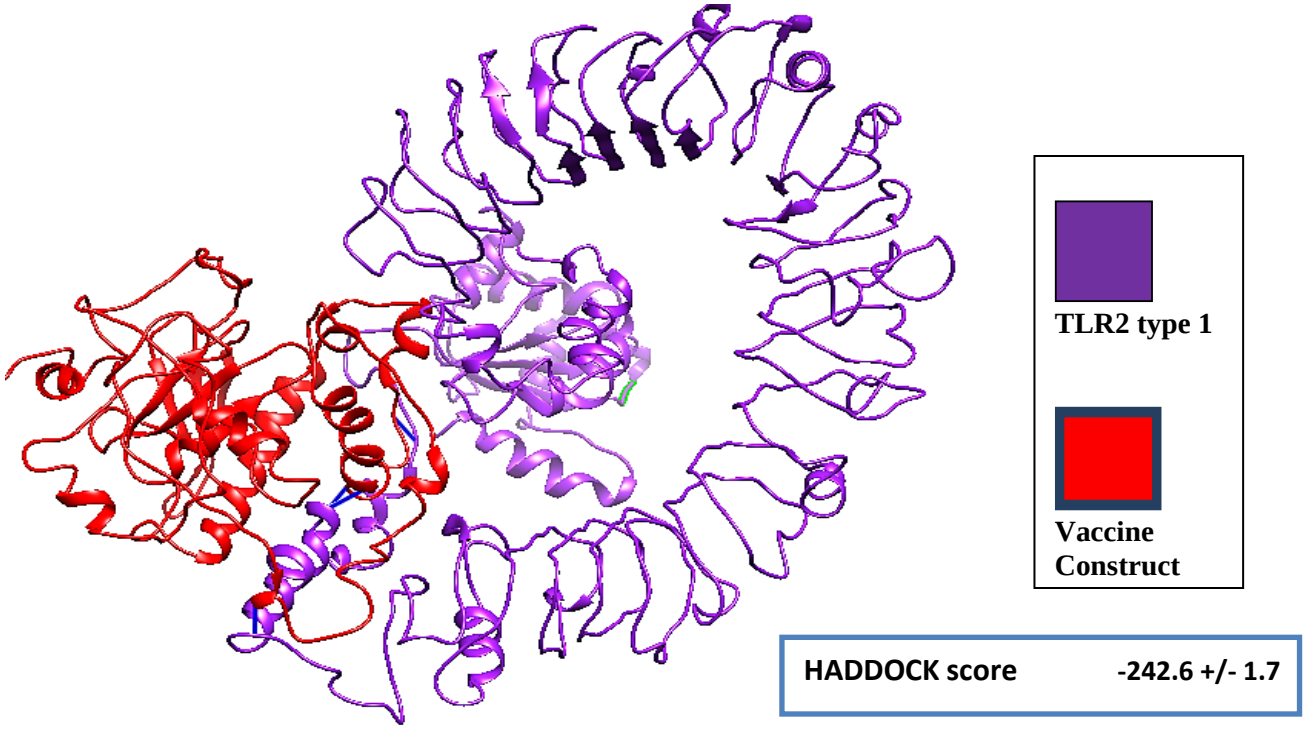


Figure 4.8 (b): Interaction of Vaccine construct with TLR2 type 1.

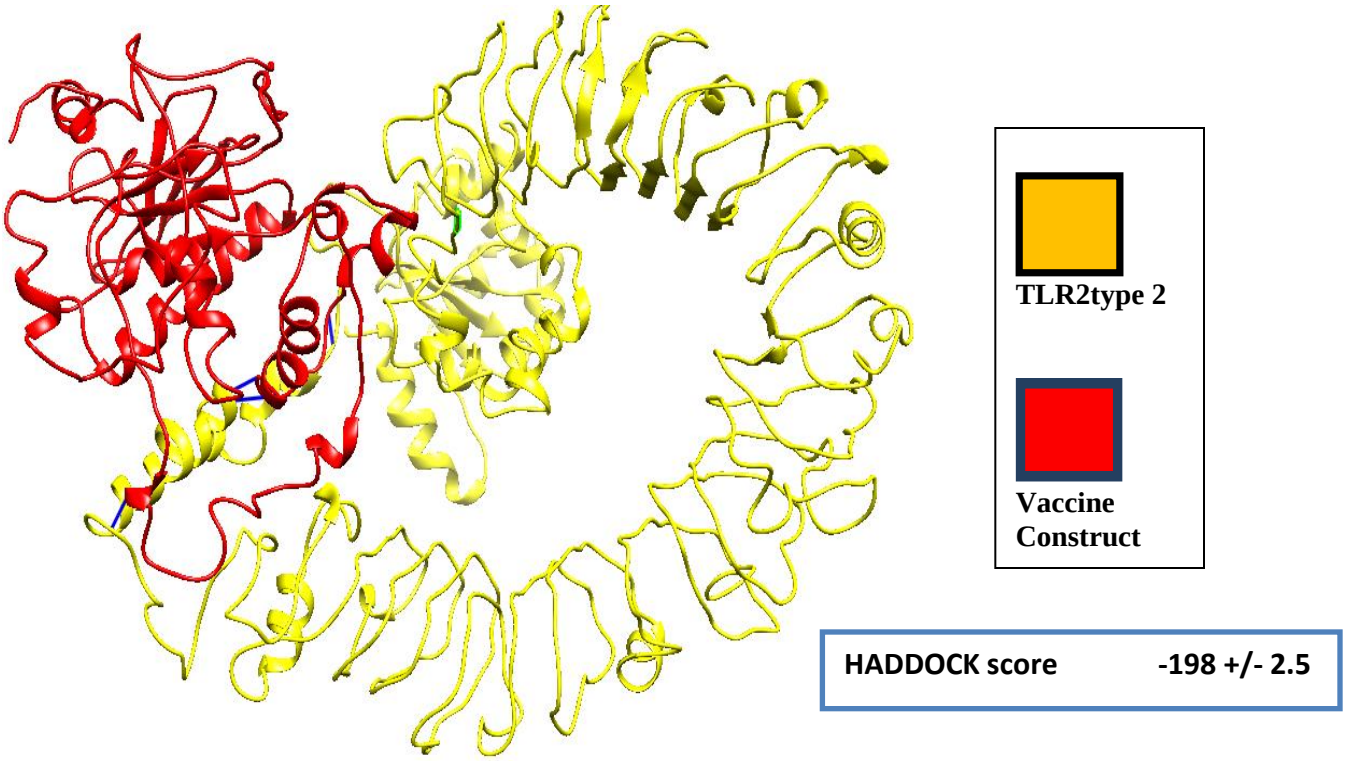
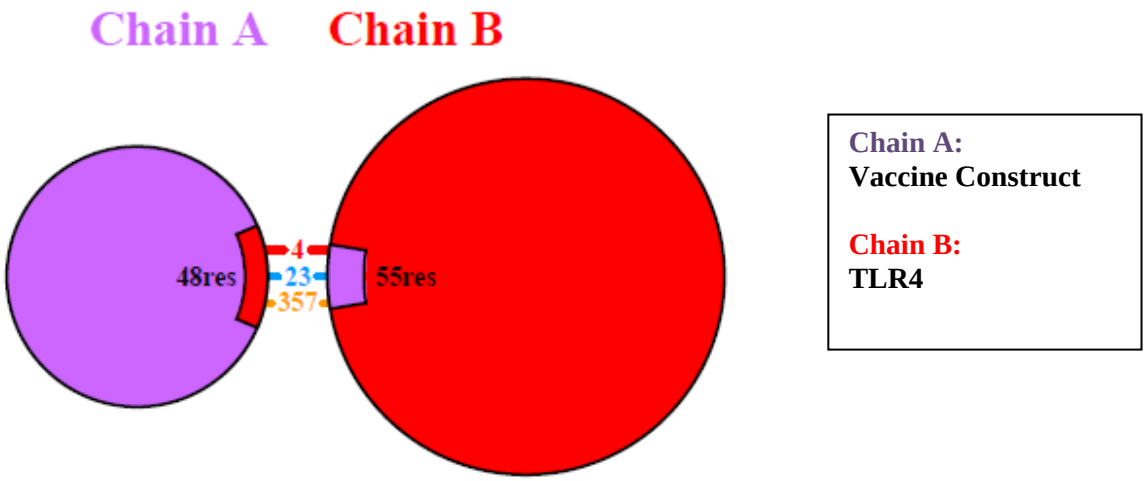


Figure 4.8 (c): Interaction of Vaccine construct with TLR2 type 2.

4.13. Docking Analysis of vaccine construct with Toll like Receptors

the detailed analysis of the selected docked cluster was further done by using PDBsum online server. This gave us a detailed insight about the residues and the intermolecular forces present within the docked cluster.

(a)

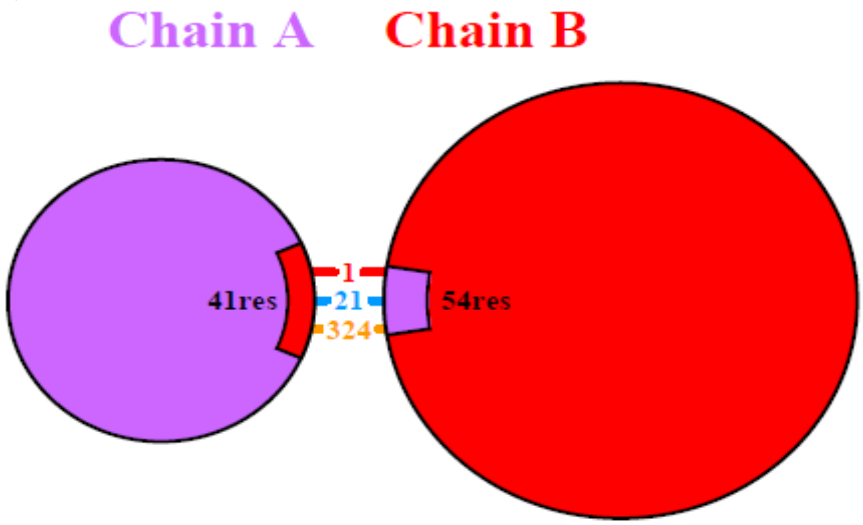


Interface statistics

Chain	No. of interface residues	Interface area (Å^2)	No. of salt bridges	No. of disulphide bonds	No. of hydrogen bonds	No. of non-bonded contacts
A	48	2487	4	-	23	357
B	55	2324				

Figure 4.9 (a): Interface analysis of docking between Vaccine construct and TLR4

(b)



Chain A:
Vaccine Construct

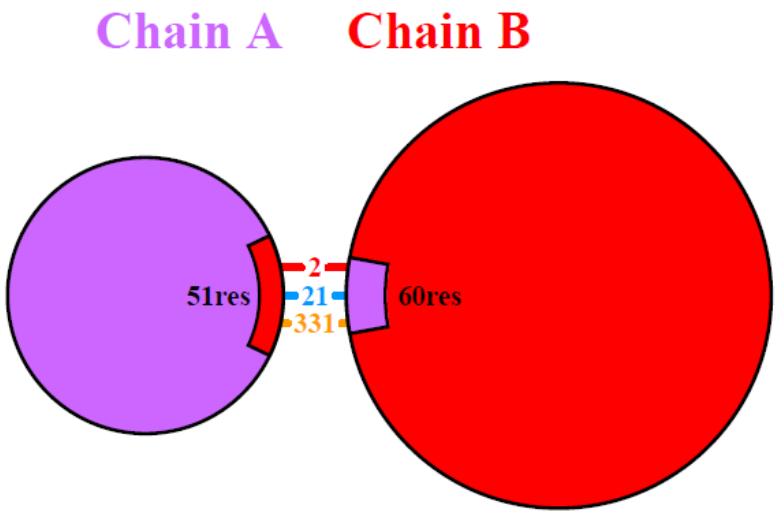
Chain B:
TLR2 type1

Interface statistics

Chain	No. of interface residues	Interface area (Å^2)	No. of salt bridges	No. of disulphide bonds	No. of hydrogen bonds	No. of non-bonded contacts
A	41	2653	1	-	21	324
B	54	2390				

Figure 4.9 (a): Interface analysis of docking between Vaccine construct and TLR2 type1

(C)



Interface statistics

Chain	No. of interface residues	Interface area (Å^2)	No. of salt bridges	No. of disulphide bonds	No. of hydrogen bonds	No. of non-bonded contacts
A	51	2788	2	-	21	331
B	60	2651				

Figure 4.9 (a): Interface analysis of docking between Vaccine construct and TLR2 type2

4.14. Analysis of Residues Interaction in Docking

The detailed analysis of the selected docked cluster was further done by using PDBsum online server. This gave us a detailed insight about the interacting residues and the intermolecular forces present within the docked cluster

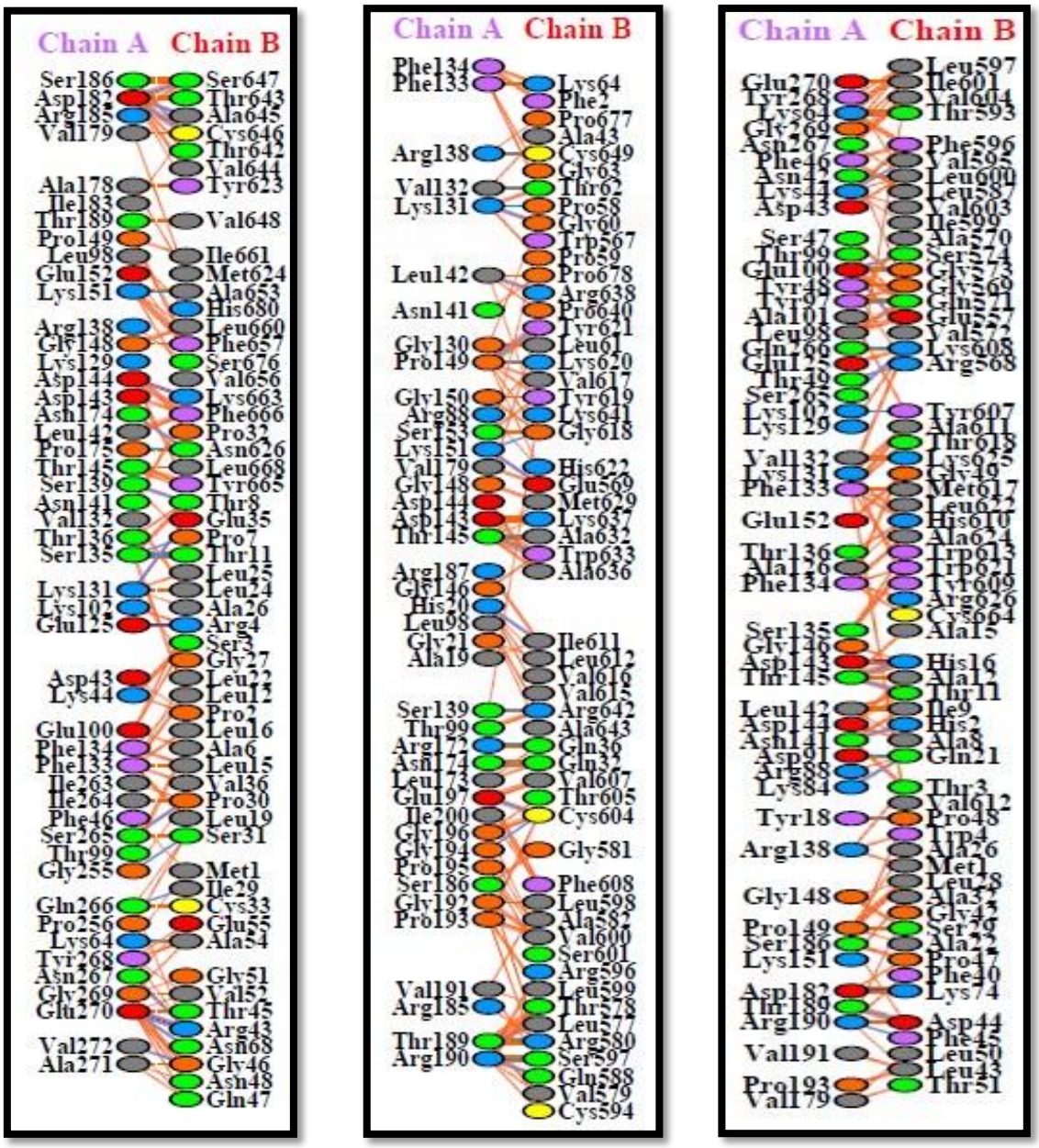


Figure 4.10: Protein-protein interaction between vaccine construct between **(a)** TLR4 **(b)** TLR2 types **1 (c)** TLR type 2

DISCUSSION

Despite the effective use of antiviral vaccine against NDV infection, the need to control the NDV spread still remains there. The outbreaks are continuously occurring within different areas of the country resulting in heavy economic loses, given that poultry industry is one of biggest industry of Pakistan. The LaSota strain, currently being in commercial live attenuated NDV vaccine, is not being successful to provide protection against NDV due to genetic variations in the field virus and other factors such as immunosuppression in broiler birds (Dimitrov et al., 2016). One of the most important factors which is responsible for the outbreaks despite the vaccination is; antigenic differences between the vaccine itself and the strains which caused the outbreak. In such cases, vaccine is unable to control the viral replication and its spread. Therefore, there always is a need of new vaccine which genotype matches the outbreak strain (Peeters & Koch 2019).

To overcome the problem of genetic variation between field virus and the live vaccine used against it, the epitopes for designing this multi-epitope vaccine were predicted from the sequences reported from the regions of Pakistan only. In this way, the vaccine designed will be closest to the genotype of NDV strains found in all over the country. A consensus sequence of both HN and F protein was obtained and that showed the most abundant residues present in the protein sequences (Schneider, 2002). This consensus sequence will also be helpful in finding the exact antigenic differences by comparing it with the reference genome available for NDV (Ranganathan, 2019). The use of consensus sequences, both genes and proteins, in designing a vaccine have been used many times in the past (Rodriguez et al., 2003; Thomson et al., 2005; Vijayachari,et al., 2015; Wan et al., 2018; Wang et al., 2019). Recently, insilico approaches have been used to develope multi-epitope vaccine from consensus sequences against COVID in 2020 (Zaheer et al., 2020; Olvera et al., 2020; Kar et al., 2020).

The strategy of reverse vaccinology used in this study is proved to be successful in developing targeted antiviral strategies. The use of modern vaccinology techniques, based on use of cell epitopes has shown strong results against various viral diseases along with cancer and malaria (Oyarzún & Kobe, 2015). Hence, the current study was focused on designing a multi-epitope vaccine, which can provide immunity against NDV infection.

As multi-epitope vaccine can provide both humoral and cellular immunities, therefore, it is a much better option than monovalent vaccines (Hasan et al., 2018).

The structural proteins HN and F of NDV were focused to design a multi-epitope vaccine because of their role in activating membrane fusion and viral entry in host cells. As the HN and F protein of NDV is involved in direct attachment with the host cell receptors and aids in the viral entry into the cell. . In 2003, Hua Li designed a recombinant protein which contained neutralizing epitopes of heamagglutinin of H392 influenza virus. These epitopes were combined and the recombinant immunogen was expressed in E.coli. ELISA results confirmed a higher titer of antibodies was induced against the H392 influenza virus as a result of using neutralizing epitope of HN (Li, 2003). In 2008, Pallavi Somvanshi, Vijay Singh, and P.K. Seth used haemmaglutinin and neuraminidase proteins of Influenza virus (Strain: H5N1) for prediction of epitopes. These epitopes were further proved beneficial in vaccine development (Somvanshi et al., 2008). In 2009, Wilfred Ndifon and his co-workers worked on the design of influenza vaccine after determining the differential neutralization efficiency of the HN epitopes. (Ndifon et al., 2009). In 2019, Mahmudul Hasan and his co-workers used the approach of reverse vaccinology for development of a subunit multi-epitope vaccine. This polyepitope vaccine was designed against avian influenza A (H7N9) virus using haemagglutinin and matrix protein 1 as the target proteins for prediction of epitopes (Hasan et al., 2019).

Emad A. Hashish and his co-workers, in 2013, used a multiepitope fusion antigen against the BVDV (bovine viral diarrhea virus) and ETEC (enterotoxigenic E.coli) (Hashish et al., 2013). In 2019, Gaafar and her co-workers designed a multiepitope vaccine against Peste des Petits Ruminants Virus or small ruminant morbillivirus. The epitopes were predicted from the fusion (F), haemagglutinin (H), matrix (M) and nucleocapsid (N) proteins of the virus using different insilico tools and softwares (Gaafar et al., 2019). In 2020, Muhammad Tahir ul Qamar and his team used the approach of reverse vaccinology in designing a multi epitope vaccine against RSV (respiratory syncytial virus) by using its fusion and glycoprotein as the target proteins for prediction of T-cell and B-cell epitopes (Tahir Ul Qamar et al., 2020).

Therefore, the epitopes predicted from these proteins will contain the necessary pathogen associated molecular patterns within them. These PAMPs are very necessary for the pathogen to be recognized by the immune system of the host (Amarante-Mendes et al., 2018). The immune system against a specific virus cannot be activated until and unless the cell receptors recognize the molecular patterns of the pathogen. In order for multi-epitope vaccine to get recognized and induce the immune response, the epitopes must be containing the necessary antigenic determinants of the viral proteins (Kar et al., 2020).

The evaluation of epitopes was done by various online tools and the prioritized epitopes were B cell, T cell and IFN gamma overlapping epitopes and were capable of inducing their respective responses. The selected epitopes were non-homologous with the chicken proteome. If we design a vaccine based on the homologous proteins from host, the vaccine will not be recognized by the host immune system as an antigen. Hence the immune system will not consider it as a threat and an immune response will not be generated (Rahman et al., 2020). The epitopes were conserved also in all strains. Selection of such conserved epitopes was necessary so that the immune response generated as a result of the vaccine, will be effective against any strain having these same conserved epitope regions in their genome (Awad Elkareem et al., 2017).

The selected B cell had overlapping T cell epitopes, possessed high immunogenic and antigenic values and non-allergen in nature, thus being the ideal candidates. These ideal candidates have a potential to generate effective, strong and targeted immune response. As the selected B cell epitopes have overlapping T cell and IFN-γ epitopes, therefore these epitopes are capable of inducing both B cell and T cell response simultaneously. Interferon gamma is also a very important cytokine with a well-known antiviral activity (Kang et al., 2018). An immunogenic adjuvant was also added which can enhance the innate and adaptive immunity. The physiochemical analyses of the vaccine projected it to be stable, hydrophilic, and acidic in nature. Various validation tools indicated the stability of the vaccine construct in nature. Urrutia-Baca and his co-workers in 2019 designed the multi-epitope vaccine against *Helicobacter pylori* by selecting the similar criteria for selection of epitopes (Urrutia-Baca et al., 2019). Dar and his co-workers also designed a

multi-epitope vaccine against *Klebsiella Pneumoniae*. They predicted, evaluated and prioritized the epitopes in similar way (Dar et al., 2019).

The epitopes in the vaccine construct were arranged on the basis of their interaction with each other. Using Haddock server, the initial pair of epitopes showing maximum binding energy were selected and joined with flexible linkers GPGPG. The arrangement (HN2-F1) was considered the best arrangement based on the refinement score of HADDOCK. Now, after joining them with linker, these were further analyzed for their binding affinity with the rest of 6 epitopes. (HN2-F1-HN3) showed the best refinement score. (HN2-F1-HN3-HN1-F3-F2-F5-F4) was the finalized construct obtained through epitope-epitope interactions. The epitopes were joined together via linkers. The selection of the linkers was done by keeping in view their properties; flexibility and their impact on the overall 3d model of the vaccine construct. The similar method was adopted by Dar and his co-workers while designing a multi-epitope vaccine against *Klebsiella pneumonae* using insilico approaches (Dar et al., 2019) and by Ikram and his team while designing a conserved multiepitope vaccine against HCV (Ikram et al., 2018).

The 3D structure prediction was carried out by ITASSER. This gave us an insight about the possible folding and three dimensional structure of our vaccine construct. The structure was further evaluated by the tools; Ramachandran plot analysis, ProSA-web- Z score and Verify 3D. All these evaluating softwares, compare the query stucture and its residues with all already available good structure in the database. Ramachandran plot has torsion angles of the amino acids in any given protein. The torsion angles involved in this plot are; psi (ψ) and phi (φ). So, by ramachandran plot, we can estimate which torsion angles are possible and permitted. Through this plot, we can know a lot about the protein structure (Ramachandran et al., 1960). Ramachandran analysis of our vaccine construct showed that 92.6% of the structure was under favorable region, 9.4% was under the allowed region, and 2.0% was observed under the disallowed region regarding it as a high quality structure. ProSA-web showed a Z-score of −6.5, which lies inside the range of acceptable scores. Online tool Verify 3D also evaluated the 3D structure as pass, signaling out its good quality (Khatoon et al., 2017).

Docking analysis was performed to analyze the binding affinity of vaccine construct with toll like receptors, TLR2-type 1, TLR2-type 2 and TLR4, of chicken. The TLRs initiate immune response against invading *pathogen by recognizing the antigens (Nie et al, 2018)*. The reason for selecting these above mentioned TLRs, were because of their location at the cell surface. The TLRs at the cell surface are directly involved in the interaction with receptors of the pathogen (Boehme & Compton, 2004). Hence, the designed vaccine was assessed for association with these receptors. The significant docking score signal out towards great interaction between vaccine construct and TLR receptors, thus enabling the vaccine to induce TLR activation which will be followed by an amplified immune response against the virus. The evaluation of vaccine construct by docking analysis with the TLR receptor has been opted many times in the past while designing a synthetic multi-epitope vaccine. (Ikram et al., 2018; Majid et al., 2019; Shey et al., 2019; Tamalika et al., 2020; Dar et al., 2019; Peele et al., 2020; Rahman et al., 2020; Jakhar et al., 2020).

CONCLUSION

Conclusion

The use of computational tools for vaccine design and its evaluation can significantly enhance the process of vaccine development and its discovery with fewer costs and less time. The in-silico designed vaccine construct has appropriate structural, physiochemical and immunological properties which can induce a strong and successful immune response against NDV. These promising results needs wet lab validation to further verify the effectiveness of this potential vaccine candidate.

Future Perspectives

In future, the experimental validation can be done to confirm the potential of this in-silico designed vaccine. Moreover, this study can be utilized in finding antigenic differences from reference genome and determination of genetic diversity among NDV strains reported globally from different regions of the world. It can also be helpful in detection of genetic determinants associated vaccine resistance in vaccinated chickens.

REFERENCES

Absalón, A. E., Cortés-Espinosa, D. V., Lucio, E., Miller, P. J., & Afonso, C. L. (2019). Epidemiology, control, and prevention of Newcastle disease in endemic regions: Latin America. *Tropical animal health and production*, *51*(5), 1033–1048.

Alexander, D. J. (1988). Newcastle Disease Diagnosis. *Newcastle Disease. Developments in Veterinary Virology,* vol 8. Springer, Boston: 147–160.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. (1990) "Basic local alignment search tool." *J. Mol. Biol.* 215:403-410.

Amarante-Mendes, G. P., Adjemian, S., Branco, L. M., Zanetti, L. C., Weinlich, R., & Bortoluci, K. R. (2018). Pattern Recognition Receptors and the Host Cell Death Molecular Machinery. *Frontiers in Immunology*, 9:2379.

Amarasinghe, G. K., Bào, Y., Basler, C. F., Bavari, S., Beer, M., Bejerman, N., Blasdell, K. R., Bochnowski, A., Briese, T., Bukreyev, A., Calisher, C. H., Chandran, K., Collins, P. L., Dietzgen, R. G., Dolnik, O., Dürrwald, R., Dye, J. M., Easton, A. J., Ebihara, H., Fang, Q., … Kuhn, J. H. (2017). Taxonomy of the order Mononegavirales: update 2017. *Archives of virology*, 162(8), 2493–2504.

Apostólico, J. de S., Lunardelli, V. A. S., Coirada, F. C., Boscardin, S. B., & Rosa, D. S. (2016). Adjuvants: Classification,Modus Operandi, and Licensing. *Journal of Immunology Research*, 1–16.

Ataman-Onal, Y., Munier, S., Ganee, A., Terrat, C., Durand, P. Y., Battail, N., Martinon, F., Le, G. R., Charles, M. H., Delair, T. and Verrier, B. (2006) Surfactant-free anionic PLA nanoparticles coated with HIV-1 p24 protein induced enhanced cellular and humeral immune responses in various animal models. *J. Control. Release* 112, 175-185.

Awad Elkareem, M. A. E., Ahmed Osman, S., Mohamed, H. A., Elrahman Hassan, H. A., Abuharaz, A. H., Abd elrahman, K. A., & Ahmed Salih, M. (2017). Prediction and Conservancy Analysis of Multiepitope Based Peptide Vaccine Against Merkel Cell Polyomavirus: An Immunoinformatics Approach. *Immunome Research*, 13(2).

Azmi, F., Ahmad Fuaad, A. A. H., Skwarczynski, M., & Toth, I. (2014). Recent progress in adjuvant discovery for peptide-based subunit vaccines. *Human vaccines & immunotherapeutics*, *10*(3), 778-796.

Beard, C.W. & Hanson, P. (1984). Newcastle disease. In: M. S. HOFSTAD (Ed.) *Diseases of Poultry*, 8th edn (Ames, Iowa State University).

Boehme, K. W., & Compton, T. (2004). Innate Sensing of Viruses by Toll-Like Receptors. *Journal of Virology*, 78(15), 7867–7873.

Brown C. & Torres A., Eds. (2008). - USAHA Foreign Animal Diseases, Seventh Edition. Committee of Foreign and Emerging Diseases of the US Animal Health Association. Boca Publications Group, Inc.

Brown, V. R., & Bevins, S. N. (2017). A review of virulent Newcastle disease viruses in the United States and the role of wild birds in viral persistence and spread. *Veterinary Research*, 48(1).

Brusic,V., Rudy,G. and Harrison, LC. (1994) Prediction of MHC binding peptides using artificial neural networks Complex system: *Mechanism of Adaptation* IOS Press/ohmsha 253-260.

Cattoli, G., Susta, L., Terregino, C., & Brown, C. (2011). Newcastle disease. *Journal of Veterinary Diagnostic Investigation*, 23(4),637–656.

Chambers, P., Millar, N. S., & Emmerson, P. T. (1986). Nucleotide Sequence of the Gene Encoding the Fusion Glycoprotein of Newcastle Disease Virus. *Journal of General Virology*, 67(12), 2685–2694.

Chong, C. S., Cao, M., Wong, W. W., Fischer, K. P., Addison, W. R., Kwon, G. S., Tyrrell, D. L., & Samuel, J. (2005). Enhancement of T helper type 1 immune responses against hepatitis B virus core antigen by PLGA nanoparticle vaccine delivery. *Journal of controlled release : official journal of the Controlled Release Society*, *102*(1), 85–99.

Cooper, C. L., Davis, H. L., Angel, J. B., Morris, M. L., Elfer, S. M., Seguin, I., Krieg, A. M., & Cameron, D. W. (2005). CPG 7909 adjuvant improves hepatitis B virus vaccine

seroprotection in antiretroviral-treated HIV-infected adults. *AIDS (London, England)*, 19(14), 1473–1479.

Cornax, I., Miller, P. J., & Afonso, C. L. (2012). Characterization of Live LaSota Vaccine Strain–Induced Protection in Chickens upon Early Challenge with a Virulent Newcastle Disease Virus of Heterologous Genotype. Avian Diseases, 56(3), 464–470.

Courtney, S. C., Susta, L., Gomez, D., Hines, N. L., Pedersen, J. C., Brown, C. C., Miller, P. J., & Afonso, C. L. (2013). Highly divergent virulent isolates of Newcastle disease virus from the Dominican Republic are members of a new genotype that may have evolved unnoticed for over 2 decades. *Journal of clinical microbiology*, *51*(2), 508–517.

Czeglédi, A., Ujvári, D., Somogyi, E., Wehmann, E., Werner, O., & Lomniczi, B. (2006). Third genome size category of avian paramyxovirus serotype 1 (Newcastle disease virus) and evolutionary implications. *Virus Research*, 120(1–2), 36–48.

Dar, H. A., Zaheer, T., Shehroz, M., Ullah, N., Naz, K., Muhammad, S. A., Zhang, T., & Ali, A. (2019). Immunoinformatics-Aided Design and Evaluation of a Potential Multi-Epitope Vaccine against *Klebsiella Pneumoniae*. *Vaccines*, *7*(3), 88.

De Vries, S. J., & Bonvin, A. M. J. J. (2011). CPORT: A Consensus Interface Predictor and Its Performance in Prediction-Driven Docking with HADDOCK. *PLoS ONE*, 6(3), e17695.

Dhanda, S. K., Vir, P., & Raghava, G. P. (2013). Designing of interferon-gamma inducing MHC class-II binders. *Biology direct*, 8, 30.

Diel, D.G., Miller, P.J., Wolf, P.C., Mickley, R.M. and Musante, A.R. (2012). Characterization of Newcastle disease viruses isolated from cormorant and gull species in the United States in 2010. *Avian Diseases*, 56: 128-133.

Dimitrov, D. S. (2004). Virus entry: molecular mechanisms and biomedical applications. *Nature Reviews Microbiology*, 2(2), 109–122.

Dimitrov, I., Naneva, L., Doytchinova, I., & Bangov, I. (2013). AllergenFP: allergenicity prediction by descriptor fingerprints. *Bioinformatics*, 30(6), 846–851.

Dimitrov, K. M., Afonso, C. L., Yu, Q., & Miller, P. J. (2017). Newcastle disease vaccines—A solved problem or a continuous challenge? *Veterinary Microbiology*, 206; 126–136.

Dominguez, C., Boelens, R., & Bonvin, A. M. J. J. (2003). HADDOCK: A Protein−Protein Docking Approach Based on Biochemical or Biophysical Information. *Journal of the American Chemical Society*, 125(7), 1731–1737.

Dortmans, J. C., Koch, G., Rottier, P. J., & Peeters, B. P. (2011). Virulence of newcastle disease virus: what is known so far? *Veterinary Research*, 42(1), 122.

Doytchinova, I. A. and Flower, D. R. (2007) Identifying candidate subunit vaccines using an alignment-independent method based on principal amino acid properties. *Vaccine*. 25:856-866.

Dutch, R. E. (2010). Entry and Fusion of Emerging Paramyxoviruses. *PLoS Pathogens*, 6(6): e1000881.

Eisenberg, D., Lüthy, R., & Bowie, J. U. (1997). VERIFY3D: assessment of protein models with three-dimensional profiles. *Methods in enzymology*, *277*, 396–404.

Estola, T., & Hovi, T. (1979). The effect of air ionization on the air-borne transmission of experimental Newcastle disease virus infections in chickens. *Epidemiology & Infection*, *83*(1), 59-67.

Etriwati, Ratih, D., Handharyani, E., & Setiyaningsih, S. (2017). Pathology and immunohistochemistry study of Newcastle disease field case in chicken in Indonesia. Veterinary world, 10(9), 1066–1071.

Farhadian, A., Dounighi, N. M., & Avadi, M. (2015). Enteric trimethyl chitosan nanoparticles containing hepatitis B surface antigen for oral delivery. *Human Vaccines & Immunotherapeutics*, 11(12), 2811–2818.

Gaafar, B., Ali, S. A., Abd-Elrahman, K. A., & Almofti, Y. A. (2019). Immunoinformatics Approach for Multiepitope Vaccine Prediction from H, M, F, and N

Proteins of Peste des Petits Ruminants Virus. *Journal of immunology research, 2019*, 6124030.

Gasteiger E., Hoogland C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., Bairoch A. (2005). Protein Identification and Analysis Tools on the ExPASy Server; (In) John M. Walker (ed): *The Proteomics Protocols Handbook, Humana Press*. pp. 571-607.

George, R. A., & Heringa, J. (2002). An analysis of protein domain linkers: their classification and role in protein folding. *Protein engineering*, 15(11), 871–879.

Gish, W. & States, D.J. (1993) "Identification of protein coding regions by database similarity search." *Nature Genet*. 3:266-272.

GOP (Government of Pakistan) (2016). Federal Bureau of Statistics, Ministry of Finance, Government of Pakistan.

Grimes, S. (2002). A Basic laboratory manual for the small-scale production and testing of 1-2 Newcastle disease vaccine. Bangkok: FAO Regional Office for Asia and the Pacific (RAP).

Hasan, M., Azim, K. F., Begum, A., Khan, N. A., Shammi, T. S., Imran, M., Chowdhury, I. M., & Urme, S. R. A. (2018). Vaccinomics strategy for developing a unique multi-epitope monovalent vaccine against Marburg marburgvirus. *BioRxiv*, 484923.

Hasan, M., Ghosh, P. P., Azim, K. F., Mukta, S., Abir, R. A., Nahar, J., & Hasan Khan, M. M. (2019). Reverse vaccinology approach to design a novel multi-epitope subunit vaccine against avian influenza A (H7N9) virus. *Microbial Pathogenesis*, 130, 19–37.

Hashish, E. A., Zhang, C., Ruan, X., Knudsen, D. E., Chase, C. C., Isaacson, R. E., Zhou, G., & Zhang, W. (2013). A multiepitope fusion antigen elicits neutralizing antibodies against enterotoxigenic Escherichia coli and homologous bovine viral diarrhea virus in vitro. *Clinical and vaccine immunology* : CVI, 20(7), 1076–1083.

HeeShin, W. H. (2014) Prediction of protein structure and interaction by GALAXY protein modeling programs. *Bio Des* 2(1):1–11.

Heo, L., Park, H., & Seok, C. (2013). GalaxyRefine: Protein structure refinement driven by side-chain repacking. *Nucleic acids research*, 41(Web Server issue), W384–W388.

Huang Z, Panda A, Elankumaran S, Govindarajan D, Rockemann DD, Samal SK (2004) The hemagglutinin-neuraminidase protein of Newcastle disease virus determines tropism and virulence. *J Virol* 78:4176–4184.

Hussain, J., Rabbani, I., Aslam, S., & Ahmad, H. A. (2015). An overview of poultry industry in Pakistan. *World's Poultry Science Journal*, 71(4), 689–700.

Ikram, A., Zaheer, T., Awan, F. M., Obaid, A., Naz, A., Hanif, R., Paracha, R. Z., Ali, A., Naveed, A. K., & Janjua, H. A. (2018). Exploring NS3/4A, NS5A and NS5B proteins to design conserved subunit multi-epitope vaccine against HCV utilizing immunoinformatics approaches. Scientific Reports, 8(1). 16107.

Jakhar, R., Kaushik, S., & Gakhar, S. K. (2020). 3CL hydrolase-based multiepitope peptide vaccine against SARS-CoV-2 using immunoinformatics. *Journal of Medical Virology*, 92(10), 2114–2123.

Jansen, S. J. T., Otten, W., & Stiggelbout, A. M. (2006). Factors affecting patients' perceptions of choice regarding adjuvant chemotherapy for breast cancer. *Breast Cancer Research and Treatment*, 99(1), 35–45.

Jiang, D., Premachandra, G. S., Johnston, C., & Hem, S. L. (2004). Structure and adsorption properties of commercial calcium phosphate adjuvant. *Vaccine*, 23(5), 693–698.

Jiao, X., Hu, Y., & Sun, L. (2010). Dissection and localization of the immunostimulating domain of Edwardsiella tarda FliC. *Vaccine*, 28(34), 5635–5640.

Kang, S., Brown, H. M., & Hwang, S. (2018). Direct Antiviral Mechanisms of Interferon-Gamma. *Immune network*, *18*(5), e33.

Kar, T., Narsaria, U., Basak, S., Deb, D., Castiglione, F., Mueller, D. M., & Srivastava, A. P. (2020). A candidate multi-epitope vaccine against SARS-CoV-2. *Scientific Reports*, 10(1).

Kavoosi, M., Creagh, A. L., Kilburn, D. G., & Haynes, C. A. (2007). Strategy for selecting and characterizing linker peptides for CBM9-tagged fusion proteins expressed in Escherichia coli. *Biotechnology and bioengineering*, *98*(3), 599–610.

Khatoon, N., Pandey, R. K., & Prajapati, V. K. (2017). Exploring Leishmania secretory proteins to design B and T cell multi-epitope subunit vaccine using immunoinformatics approach. *Scientific Reports*, 7(1).

Ko, J., Park, H., Heo, L., & Seok, C. (2012). GalaxyWEB server for protein structure prediction and refinement. *Nucleic acids research*, *40*(Web Server issue), W294–W297.

Kool, M., Fierens, K., & Lambrecht, B. N. (2012). Alum adjuvant: some of the tricks of the oldest adjuvant. *Journal of Medical Microbiology*, 61(7), 927–934.

Krammer, F., Palese, P., & Steel, J. (2014). Advances in Universal Influenza Virus Vaccine Design and Antibody Mediated Therapies Based on Conserved Regions of the Hemagglutinin. *Current Topics in Microbiology and Immunology* (pp. 301–321). Springer International Publishing.

Kumar Jaiswal, A., Tiwari, S., Jamal, S.B., Barh, D., Azevedo, V., Soares, S.C. (2017) An In Silico Identification of Common Putative Vaccine Candidates against *Treponema pallidum*: A Reverse Vaccinology and Subtractive Genomics Based Approach. *Int. J. Mol. Sci*,18(2):402.

Lancaster, J. (1976). A History of Newcastle Disease with Comments on its Economic Effects. *World's Poultry Science Journal*, *32*(2), 167-175.

Laskowski, R. A. (2004). PDBsum more: new summaries and analyses of the known 3D structures of proteins and nucleic acids. *Nucleic Acids Research*, 33(Database issue), D266–D268.

Lei, Y., Zhao, F., Shao, J., Li, Y., Li, S., Chang, H., & Zhang, Y. (2019). Application of built-in adjuvants for epitope-based vaccines. *PeerJ*, *6*, e6185.

Leroux-Roels, I., Borkowski, A., Vanwolleghem, T., Dramé, M., Clement, F., Hons, E., Devaster, J.-M., & Leroux-Roels, G. (2007). Antigen sparing and cross-reactive

immunity with an adjuvanted rH5N1 prototype pandemic influenza vaccine: a randomised controlled trial. *The Lancet*, 370(9587), 580–589.

Li, H. (2003). Recombinant protein comprising multi-neutralizing epitopes induced high titer of antibodies against Influenza A virus. Immunobiology, 207(5), 305–313.

Liang, S., Zheng, D., Zhang, C., & Zacharias, M. (2009). Prediction of antigenic epitopes on protein surfaces by consensus scoring. *BMC Bioinformatics*, 10(1), 302.

Liu, J., Zhu, J., Xu, H., Li, J., Hu, Z., Hu, S., And, X. W., & Liu, X. (2017). Effects of the HN Antigenic Difference between the Vaccine Strain and the Challenge Strain of Newcastle Disease Virus on Virus Shedding and Transmission. *Viruses*, 9(8), 225.

L′vov, D., Shchelkanov, M., Alkhovsky, S. & Deryabin, P. (2015). *Zoonotic viruses of Northern Eurasia : taxonomy and ecology*. Amsterdam: Academic Press. Thomas, D. R., & Walmsley, A. M. (2018). Plant-Made Veterinary Vaccines for Newcastle Disease Virus. *Prospects of Plant-Based Vaccines in Veterinary Medicine*; pp. 149–167. Springer International Publishing.

Manoharan, V. K., Khattar, S. K., LaBranche, C. C., Montefiori, D. C., & Samal, S. K. (2018). Modified Newcastle Disease virus as an improved vaccine vector against Simian Immunodeficiency virus. *Scientific Reports*, 8(1): 8952.

Marchetti, M., Rossi, M., Giannelli, V., Giuliani, M. M., Pizza, M., Censini, S., Covacci, A., Massari, P., Pagliaccia, C., Manetti, R., Telford, J. L., Douce, G., Dougan, G., Rappuoli, R., & Ghiara, P. (1998). Protection against Helicobacter pylori infection in mice by intragastric vaccination with H. pylori antigens is achieved using a non-toxic mutant of E. coli heat-labile enterotoxin (LT) as adjuvant. *Vaccine*, 16(1), 33–37.

McEwen, J., Levi, R., Horwitz, R. J., & Arnon, R. (1992). Synthetic recombinant vaccine expressing influenza haemagglutinin epitope in Salmonella flagellin leads to partial protection in mice. *Vaccine*, 10(6), 405–411.

Mehrabi, M., Montazeri , H., Mohammadpour Dounighi, N., Rashti, A., Vakili-Ghartavol, R. (2018) . Chitosan-based Nanoparticles in Mucosal Vaccine Delivery. *Archives of Razi Institute,* 73(3): 165-176.

Mishra, N., Gupta, P. N., Mahor, S., Khatri, K., Goyal, A. K., & Vyas, S. P. (2007). Liposomes as adjuvant for combination vaccines. *Indian journal of experimental biology, 45*(3), 237–241.

Mogensen T. H. (2009). Pathogen recognition and inflammatory signaling in innate immune defenses. *Clinical microbiology reviews*, *22*(2), 240–273.

Mohammadpourdounighi, N., Behfar, A., Ezabadi, A., Zolfagharian, H., & Heydari, M. (2010). Preparation of chitosan nanoparticles containing Naja naja oxiana snake venom. *Nanomedicine: Nanotechnology, Biology and Medicine*, *6*(1), 137-143.

Mohammadpourdounighi, N., Dmavandi, M., Zolfagharian, H., Dehgahi, S. M.. (2012). Preparing and Characterizing Chitosan Nanoparticles Containing Hemiscorpius lepturus Scorpion Venom as an Antigen Delivery System. *Archives of Razi Institute* 67(2):145-153.

MohammadpourDounighi, N., Javadi, M., Zolfagharian, H., Awadi, M.R. (2016). Preparation and characterization of Agkistrodon Halys venom entrapped chitosan nanoparticles: Novel and advanced antigen delivery and adjuvant system. *Journal of Kerman University of Medical Sciences,* 23(5):618-630.

Moyle & Toth (2013) Moyle PM, Toth I. Modern subunit vaccines: development, components, and research opportunities. *ChemMedChem*, 8(3):360–376.

Nagai Y, Klenk HD (1977) Activation of precursors to both glycoproteins of Newcastle disease virus by proteolytic cleavage. *Virology* 77:125–134.

Ndifon, W., Wingreen, N. S., & Levin, S. A. (2009). Differential neutralization efficiency of hemagglutinin epitopes, antibody interference, and the design of influenza vaccines. *Proceedings of the National Academy of Sciences*, 106(21), 8701–8706.

Nie, L., Cai, S.-Y., Shao, J.-Z., & Chen, J. (2018). Toll-Like Receptors, Associated Biological Roles, and Signaling Networks in Non-Mammals. *Frontiers in Immunology*, 9: 1523.

Nimbalkar, V.V., Kadu, U.E., Shelke, R.P.,Gaikwad, P.M. (2018) A Review on: Overview of Immune System, Immunomodulation and Experimental Animal Models for Evaluation of Potential Immunomodulators. *Ijppr.Human,* 2 (4): 151-164.

Okonechnikov, K., Golosova, O., Fursov, M., & UGENE team (2012). Unipro UGENE: a unified bioinformatics toolkit. *Bioinformatics (Oxford, England), 28(8),* 1166–1167.

O'Leary, N. A., Wright, M. W., Brister, J. R., Ciufo, S., Haddad, D., McVeigh, R., Rajput, B., Robbertse, B., Smith-White, B., Ako-Adjei, D., Astashyn, A., Badretdin, A., Bao, Y., Blinkova, O., Brover, V., Chetvernin, V., Choi, J., Cox, E., Ermolaeva, O., Farrell, C. M., … Pruitt, K. D. (2016). Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic acids research*, *44*(D1), D733–D745.

Olvera, A., Noguera-Julian, M., Kilpelainen, A., Romero-Martín, L., Prado, J. G., & Brander, C. (2020). SARS-CoV-2 Consensus-Sequence and Matching Overlapping Peptides Design for COVID19 Immune Studies and Vaccine Development. *Vaccines*, 8(3), 444.

Oyarzún, P. & Kobe, B. (2015). Recombinant and epitope-based vaccines on the road to the market and implications for vaccine design and production. *Human Vaccines & Immunotherapeutics,* 12(3), 763–767.

Peeters, B., & Koch, G. (2019). Newcastle Disease Virus. *Reference Module in Life Sciences.* Elsevier.

Phale, S. (2018). Newcastle Disease Virus: Structural and Molecular Basis of Pathogenicity. *Medicinal Chemistry*, 08:08.

Rahman, N., Ajmal, A., Ali, F., & Rastrelli, L. (2020). Core proteome mediated therapeutic target mining and multi-epitope vaccine design for Helicobacter pylori. *Genomics*, 112(5), 3473–3483.

Rahman, N., Ali, F., Basharat, Z., Shehroz, M., Khan, M. K., Jeandet, P., Nepovimova, E., Kuca, K., & Khan, H. (2020). Vaccine Design from the Ensemble of Surface Glycoprotein Epitopes of SARS-CoV-2: An Immunoinformatics Approach. Vaccines, 8(3), 423.

Ramachandran, G. N., & Sasisekharan, V. (1968). Conformation of Polypeptides and Proteins. *Advances in Protein Chemistry*, Volume 23 (pp. 283–437).

Ranganathan, S., Gribskov, M., Nakai, K. & Schönbach, C. (2019). Encyclopedia of bioinformatics and computational biology. Amsterdam, Netherlands: Elsevier.

Rao, P. L., Gandham, R. K., & Subbiah, M. (2020). Molecular evolution and genetic variations of V and W proteins derived by RNA editing in Avian Paramyxoviruses. *Scientific Reports*, 10(1).

Reed, S. G., Bertholet, S., Coler, R. N., & Friede, M. (2009). New horizons in adjuvants for vaccine development. *Trends in Immunology*, 30(1), 23–32.

Rehan, M., Aslam, A., Khan, M.-R., Abid, M., Hussain, S., Umber, J., Anjum, A., & Hussain, A. (2019). Potential Economic Impact of Newcastle Disease Virus Isolated from Wild Birds on Commercial Poultry Industry of Pakistan: A Review. *Hosts and Viruses*, 6(1):1-15.

Roberts, J. R., Souillard, R., & Bertin, J. (2011). Avian diseases which affect egg production and quality. *Improving the Safety and Quality of Eggs and Egg Products*, 376–393.

Rodriguez, L. L., Barrera, J., Kramer, E., Lubroth, J., Brown, F., & Golde, W. T. (2003). A synthetic peptide containing the consensus sequence of the G–H loop region of foot-and-mouth disease virus type-O VP1 and a promiscuous T-helper epitope induces

peptide-specific antibodies but fails to protect cattle against viral challenge. Vaccine, 21(25–26), 3751–3756.

Roy, A., Kucukural, A., Zhang, Y. (2010) I-TASSER: a unified platform for automated protein structure and function prediction. Nature Protocols, 5: 725-738.

Sadiq M (2004). Pakistan poultry sector still on an upward swing. *World Poult. J*. 20: 10–11.

Saha, S., & Raghava, G. P. (2006). Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. *Proteins, 65(1),* 40–48.

Sastre, P., Ruiz, T., Schildgen, O., Schildgen, V., Vela, C., & Rueda, P. (2012). Seroprevalence of human respiratory syncytial virus and human metapneumovirus in healthy population analyzed by recombinant fusion protein-based enzyme linked immunosorbent assay. Virology Journal, 9(1): 30.

Satyanarayana, S., Krishna, M., Pavan Kumar, P., & Jeereddy, S. (2018). *In silico* structural homology modeling of nif A protein of rhizobial strains in selective legume plants. *Journal, genetic engineering & biotechnology*, *16*(2), 731–737.

Schneider T. D. (2002). Consensus sequence Zen. Applied bioinformatics, 1(3), 111–119.

Schwarz, T. F., & Leo, O. (2008). Immune response to human papillomavirus after prophylactic vaccination with AS04-adjuvanted HPV-16/18 vaccine: Improving upon nature. *Gynecologic Oncology*, 110(3), S1–S10.

Sette, A., & Fikes, J. (2003). Epitope-based vaccines: an update on epitope identification, vaccine design and delivery. *Current Opinion in Immunology*, 15(4), 461–470.

Sette, A., Rappuoli, R. (2010) Reverse vaccinology: developing vaccines in the era of genomics. *Immunity*; 33(4):530-41.

Sjölander A, Drane D, Maraskovsky E, Scheerlinck J, Suhrbier A, Tennent J, Pearse M. (2001). Immune responses to ISCOM formulations in animal and primate models. *Vaccine*. ;19(17-19):2661–2665.

Soema, P. C., Kompier, R., Amorij, J.-P., & Kersten, G. F. A. (2015). Current and next generation influenza vaccines: Formulation and production strategies. *European Journal of Pharmaceutics and Biopharmaceutics*, 94, 251–263.

Somvanshi, P., Singh, V., & Seth, P. K. (2008). Prediction of Epitopes in Hemagglutinin and Neuraminidase Proteins of Influenza A Virus H5N1 Strain: A Clue for Diagnostic and Vaccine Development. OMICS: A Journal of Integrative Biology, 12(1), 61–69.

Spradbrow, P.B., Samuel, J.L. and Ibrahim, A.L.(1988). Serological response of chickens to oral vaccination with Newcastle disease virus. *J. Vet. Microbiol.*,16: 255-262.

Su, Q., Li, Y., Zhang, Y., Zhang, Z., Meng, F., Cui, Z., Chang, S., & Zhao, P. (2018). Newcastle disease virus-attenuated vaccine LaSota played a key role in the pathogenicity of contaminated exogenous virus. Veterinary Research, 49(1):80.

Sun, H. X., Xie, Y., & Ye, Y. P. (2009). Advances in saponin-based adjuvants. *Vaccine*, *27*(12), 1787–1796.

Tahir Ul Qamar, M., Shokat, Z., Muneer, I., Ashfaq, U. A., Javed, H., Anwar, F., Bari, A., Zahid, B., & Saari, N. (2020). Multiepitope-Based Subunit Vaccine Design and Evaluation against Respiratory Syncytial Virus Using Reverse Vaccinology Approach. *Vaccines*, 8(2), 288.

Thomson, S. A., Jaramillo, A. B., Shoobridge, M., Dunstan, K. J., Everett, B., Ranasinghe, C., Kent, S. J., Gao, K., Medveckzy, J., Ffrench, R. A., & Ramshaw, I. A. (2005). Development of a synthetic consensus sequence scrambled antigen HIV-1 vaccine designed for global use. *Vaccine*, 23(38), 4647–4657.

Tran, N. T., Jakovlić, I., & Wang, W.M. (2015). In silico characterisation, homology modelling and structure-based functional annotation of blunt snout bream (Megalobrama amblycephala) Hsp70 and Hsc70 proteins. *Journal of Animal Science and Technology*, 57(1):44.

Tu, H. A., Deeks, S. L., Morris, S. K., Strifler, L., Crowcroft, N., Jamieson, F. B., Kwong, J. C., Coyte, P. C., Krahn, M., & Sander, B. (2014). Economic evaluation of meningococcal serogroup B childhood vaccination in Ontario, Canada. *Vaccine*, *32*(42), 5436–5446.

Urrutia-Baca, V. H., Gomez-flores, R., De La Garza-Ramos, M. A., Tamez-guerra, P., Lucio-sauceda, D. G., & Rodríguez-padilla, M. C. (2019). Immunoinformatics Approach to Design a Novel Epitope-Based Oral Vaccine Against Helicobacter pylori. *Journal of Computational Biology*, 26(10), 1177–1190.

Vijayachari, P., Vedhagiri, K., Mallilankaraman, K., Mathur, P., Sardesai, N., Weiner, D., Ugen, K., & Muthumani, K. (2015). Immunogenicity of a novel enhanced consensus DNA vaccine encoding the leptospiral protein LipL45. Human Vaccines & *Immunotherapeutics*, 11(8), 1945–1953.

Vita, R., Mahajan, S., Overton, J. A., Dhanda, S. K., Martini, S., Cantrell, J. R., Wheeler, D. K., Sette, A., & Peters, B. (2019). The Immune Epitope Database (IEDB): 2018 update. *Nucleic acids research*, *47*(D1), D339–D343.

Wakamatsu, N., King, D. J., Kapczynski, D. R., Seal, B. S., & Brown, C. C. (2006). Experimental pathogenesis for chickens, turkeys, and pigeons of exotic Newcastle disease virus from an outbreak in California during 2002-2003. *Veterinary pathology*, *43*(6), 925-933.

Wan, Y., Kang, G., Sreenivasan, C., Daharsh, L., Zhang, J., Fan, W., Wang, D., Moriyama, H., Li, F., & Li, Q. (2018). A DNA Vaccine Expressing Consensus Hemagglutinin-Esterase Fusion Protein Protected Guinea Pigs from Infection by Two Lineages of Influenza D Virus. *Journal of Virology*, 92(11).

Wang, W., Xia, M., Chen, J., Deng, F., Yuan, R., Zhang, X., & Shen, F. (2016). Data set for phylogenetic tree and RAMPAGE Ramachandran plot analysis of SODs in Gossypium raimondii and G. arboreum. Data in Brief, 9, 345–348.

Wen G, Partridge M, Li B, Hong M, Liao W, Cheng S, Zhao Y, Calaf G, Liu T, Zhou J, Zhang Z, Hei T (2011) TGFBI expression reduces *in vitro* and *in vivo* metastatic potential of lung and breast tumor cells. *Cancer Lett* 308: 23–32.

Wiederstein, M., & Sippl, M. J. (2007). ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic acids research*, *35*(Web Server issue), W407–W410.

Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., Zhang, Y. (2015) The I-TASSER Suite: Protein structure and function prediction. Nature Methods, 12: 7-8.

Yang, J., Zhang, Y. (2015). I-TASSER server: new development for protein structure and function predictions. Nucleic Acids Research, 43: W174-W181.

Yao, B., Zhang, L., Liang, S., & Zhang, C. (2012). SVMTriP: A Method to Predict Antigenic Epitopes Using Support Vector Machine to Integrate Tri-Peptide Similarity and Propensity. *PLoS ONE*, 7(9), e45152.

Zheng, Z., Liang, W., Wang, D., Schroder, P. M., Ju, W., Wu, L., Zheng, Z., Shang, Y., Guo, Z., & He, X. (2014). Adjuvant chemotherapy for patients with primary hepatocellular carcinoma: A meta-analysis. *International Journal of Cancer*, 136(6), E751–E759.