MULTI-EPITOPE PEPTIDE VACCINE PREDICTION AGAINST NEWCASTLE DISEASE VIRUS USING IMMUNO-INFORMATICS APPROACHES

N. PUTRI¹, I. WULANDARI¹, R. ERNAWATI² & F. ABDUL RANTAM²,³

¹Faculty of Veterinary Medicine, Universitas Airlangga, Surabaya, Indonesia; ²Laboratory of Virology and Immunology, Department of Microbiology, Faculty of Veterinary Medicine, Universitas Airlangga, Surabaya, Indonesia; ³Stem Cell Research and Development Center, Universitas Airlangga, Surabaya, Indonesia

Summary


Newcastle disease is one of the most critical disease in poultry and wild birds, largely due to its high morbidity and mortality, as well as its worldwide distribution and threat of considerable economic losses to avian industries caused by Newcastle disease virus (NDV). The NDV can cause clinical signs varying from subclinical infections to 100% mortality, depending on the susceptibility of the host and the virulence of the virus. The virus is classified into velogenic (viscerotropic velogenic and neurotropic velogenic), mesogenic and lentogenic. The objectives of this study was to design a peptide vaccine using immunoinformatics approaches. In total, 12 NDV fusion proteins retrieved from NCBI database were aligned to determine the conservancy and candidate epitopes were analysed by predictions tools from Immune Epitope Database. Then the 3D structure of the conserved region was modelled using the Swiss Model and aligned using PyMol software. Two epitopes were predicted as a peptide vaccine for B cell (DKAVNYTSSQT and NMPKDKEACAKAPEA). This is a preliminary study of designing an epitope-based peptide vaccine against NDV, and we recommend further study to identify the interaction between these peptides with T cells and antibodies.

Key words: epitope, Immune Epitope Database (IEDB), Newcastle disease virus (NDV), peptide vaccine

INTRODUCTION

Newcastle disease has been regarded as an important poultry disease and a major cause of economic loss in the poultry industry (Sadiq & Mohammed, 2017). The causative agent is Newcastle disease virus (NDV), which is also known as Avian Paramyxovirus-1 (APMV-1) of the genus Avulavirus belonging to the family of Paramyxovirus serotypes (Dimitrov et al., 2016). The first appearance of NDV in Jakarta was in 1926 reported by Kraneveld (Kumar et al., 2011).
The genome of NDV consists of six transcriptional units to encode (Miller & Koch, 2013): nucleocapsid, phosphoprotein, matrix, fusion, haemagglutinin–neuraminidase, and polymerase. Fusion protein on NDV directs membrane fusion between the viral and the cellular membranes (Morrison, 2003). Based on the mean times in which they kill inoculated chicken embryos and their virulences for day-old chickens, NDV strains can be categorised into four pathotypes, i.e., nonvirulent, lentogenic (low virulence), mesogenic (intermediate), and velogenic (highly virulent) (Alexander & Gough, 2003). The cleavage site of fusion protein is required to initiate infection and considered to determine NDV virulence (Samal, 2012). Generally, the major means of prevention against the highly virulent ND is by vaccination, which is achievable with the low pathogenic genotypes attributed to the serological similarity between the NDV genotypes (OIE, 2013; Abdul Masum et al., 2014).

The aim of this study was to design a vaccine for NDV using peptide of its fusion protein as an immunogen to stimulate protective immune response. In silico prediction of epitopes of appropriate protein residues would help in production of peptide vaccine with powerful immunogenic and minimal allergenic effect. This is the first study conducted to design a peptide vaccine against Newcastle disease virus using an immunoinformatics approach.

**MATERIALS AND METHODS**

**Protein sequence retrieval**

A total of 12 NDV strain fusion proteins were retrieved from NCBI (http://www.ncbi.nlm.nih.gov/protein/?term=newcastle disease virus+fusionprotein) database in January 2019. These sequences were retrieved from 12 strains from different countries in Indonesia. Retrieved fusion proteins and their accession numbers and area of collection are listed in Table 1.

**Phylogenetic and alignment study**

The retrieved sequences were used in phylogenetic and alignment study to determine the ancestor of each sequence and the conservancy using MEGA6 software.

**Conserved region identification**

The retrieved sequences were aligned to obtain conserved regions using multiple sequence alignment. Sequences were aligned with the ClustalW as implemented in Table 1.

**Table 1.** Virus strains retrieved and their accession numbers and area of collection

<table>
<thead>
<tr>
<th>Number</th>
<th>Accession Number</th>
<th>Date of collection</th>
<th>Country</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>HQ697255.1</td>
<td>2010</td>
<td>Sukorejo</td>
</tr>
<tr>
<td>2</td>
<td>HQ69725</td>
<td>2010</td>
<td>Banjarmasin</td>
</tr>
<tr>
<td>3</td>
<td>HQ697257.1</td>
<td>2010</td>
<td>Gianyar</td>
</tr>
<tr>
<td>4</td>
<td>HQ697258.1</td>
<td>2010</td>
<td>Sragen</td>
</tr>
<tr>
<td>5</td>
<td>HQ697260.1</td>
<td>2010</td>
<td>Kudus</td>
</tr>
<tr>
<td>6</td>
<td>HQ697261.1</td>
<td>2010</td>
<td>Bali</td>
</tr>
<tr>
<td>7</td>
<td>HQ697259.1</td>
<td>2010</td>
<td>Kudus</td>
</tr>
<tr>
<td>8</td>
<td>HQ697256.1</td>
<td>2009</td>
<td>Makasar</td>
</tr>
<tr>
<td>9</td>
<td>AB605247.1</td>
<td>2007</td>
<td>Bali</td>
</tr>
<tr>
<td>10</td>
<td>JX393313.1</td>
<td>1997</td>
<td>Jakarta</td>
</tr>
<tr>
<td>11</td>
<td>AY56298</td>
<td>1990</td>
<td>Jakarta</td>
</tr>
<tr>
<td>12</td>
<td>KF767106.1</td>
<td>1976</td>
<td>Jakarta</td>
</tr>
</tbody>
</table>
in the Biological Sequence Alignment Editor (BioEdit) version 8.0 (Chevenet et al., 2006) for finding the conserved regions among NDV. Later on, the candidate epitopes were analysed by predictions tools from Immune Epitope Database IEDB analysis resource (Hall, 1999). Then the 3D structure of the conserved region was modelled using the Swiss Model (www.swissmodel.expasy.org) (Biehnert et al., 2018) and this 3D structure was aligned using PyMol software with the template obtained to confirm the conserved region of NDV fusion protein.

**RESULTS**

The phylogenetic trees revealed that the strains of NDV collected from Gianyar, Sragen, Kudus and Banjarmasin in 2010 could be the same. Strains from Makasar 2009, Sukorejo 2010 and Bali 2010 could be the same and strains from Jakarta 1976 could be the same as 1977 (Fig. 1).

The multiple alignment results for the amino acid sequence of the protein coding gene F Newcastle disease virus was presented (Table 2). After obtaining the conserved region, the protein was modelled using homology modelling method. The result showed that the conserved region was 95.51% similar to the 3maw.1 template identified as a NDV fusion protein (Fig. 2A).

The B cell epitope analysis

B cell epitope is accessible and antigenic (Vita et al., 2015). To predict the B cell epitope, we used the methods provided by immune epitope database (http://toolsseidb.org/bcell/) with default threshold BepiPred value of 0.35 (Anayet et al., 2013); Emini surface accessibility (Emini et al., 1985) with threshold 1.000; Kolaskar-Tongaonkar antigenicity (Kolaskar & Tongaonkar, 1990) with threshold 1.015.
### Table 2. Multiple sequence alignment using BioEdit tool. Dots show the conservancy between sequences.

<table>
<thead>
<tr>
<th>Position</th>
<th>NDV/Indonesia</th>
<th>NDV/JKT/1997/JX393313.1</th>
<th>NDV/Makasar</th>
<th>NDV/Kudus</th>
<th>NDV/Sragen</th>
<th>NDV/Bali</th>
<th>NDV/Banjarmasin</th>
<th>NDV/Gianyar</th>
<th>NDV/Sukorejo</th>
<th>NDV/JKT/1976/KF767106.1</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>K</td>
<td>V</td>
<td>I</td>
<td>V</td>
<td>K</td>
<td>V</td>
<td>K</td>
<td>V</td>
<td>K</td>
<td>K</td>
</tr>
<tr>
<td>2</td>
<td>R</td>
<td>D</td>
<td>L</td>
<td>D</td>
<td>R</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>R</td>
<td>N</td>
</tr>
<tr>
<td>3</td>
<td>S</td>
<td>H</td>
<td>N</td>
<td>N</td>
<td>S</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>4</td>
<td>G</td>
<td>S</td>
<td>M</td>
<td>M</td>
<td>G</td>
<td>S</td>
<td>M</td>
<td>M</td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td>5</td>
<td>D</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>D</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>D</td>
<td>D</td>
</tr>
<tr>
<td>6</td>
<td>E</td>
<td>Q</td>
<td>Q</td>
<td>Q</td>
<td>E</td>
<td>Q</td>
<td>Q</td>
<td>Q</td>
<td>E</td>
<td>E</td>
</tr>
<tr>
<td>7</td>
<td>L</td>
<td>V</td>
<td>V</td>
<td>V</td>
<td>L</td>
<td>V</td>
<td>V</td>
<td>V</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>8</td>
<td>D</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>D</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>D</td>
<td>D</td>
</tr>
<tr>
<td>9</td>
<td>G</td>
<td>Q</td>
<td>Q</td>
<td>Q</td>
<td>G</td>
<td>Q</td>
<td>Q</td>
<td>Q</td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td>10</td>
<td>E</td>
<td>Q</td>
<td>Q</td>
<td>Q</td>
<td>E</td>
<td>Q</td>
<td>Q</td>
<td>Q</td>
<td>E</td>
<td>E</td>
</tr>
</tbody>
</table>

Multi-epitope peptide vaccine prediction against Newcastle disease virus using immuno-informatics...
The surface accessibility scores of the NDV fusion protein were 1.000 with a maximum and minimum values 7.786 and 0.095 respectively (Fig. 4). All values equal to or greater than the threshold were located on the surface. The default threshold of the Kolaskar and Tongaonkar antigenicity in this study was 1.046 with a maximum and minimum values were 1.208 and 0.879 (Fig. 5).

**DISCUSSION**

Development of an epitope based peptide vaccine has become a key motif in case of viral vaccine preparation with the advancement of sequencing technologies and the huge disclosure of protein data. The first outbreaks of ND in Java, Indonesia, and Newcastle-upon-Type, England, were reported during the mid-1920s (Kra-
Multi-epitope peptide vaccine prediction against Newcastle disease virus using immuno-informatics...

neveld, 1926; Doyle, 1927). Since then, ND has been reported everywhere. Until now, there is no single area in Indonesia free from this disease despite the vaccination program, as part of prevention, that has been done since 1950 (Kusumaningsih & Bahri, 2005).

In our study, we predicted the B cell epitopes of the conserved region of the NDV fusion protein based on the BepiPred method, Emini surface accessibility, Kolaskar Tongaonkar antigenicity on IEDB. These methods were used to predict specific areas in proteins that bind to the B cell receptor, and must be on the surface and immunogenic. Bepipred linear epitope prediction tools is the programme based on a Hidden Markov model, the best single method for predicting linear epitope and a trend scale which the positive predictions are characterised by E as B cell linear epitope (Larsen et al., 2006). The Emini surface accessibility scale describes that every peptide having a surface probability value of 1.000 is located on the cell surface (Emini et al., 1985). The Kolaskar and Tongaonkar antigenicity has 75% accuracy (Kolaskar & Tongaonkar, 1990). This method works based on the physicochemical properties of amino acid residues and the known tendency frequency as an experimental epitope.

The results of the epitope predictions indicated that there were potential peptide

![Fig. 4. Emini surface accessibility prediction. Yellow areas above threshold (red line) are proposed to be a part of B cell epitope; green areas below the threshold are not.](image)

![Fig. 5. Kolaskar and tongaonkar antigenicity prediction. Yellow areas above threshold (red line) are proposed to be a part of B cell epitope; green areas below the threshold are not.](image)
sequences as the B cell epitopes. We predicted that the peptide sequences from 47–58, 68–83 (DKAVNVYTSSQT and NMPKDKKEACAKAPEA) amino acids are capable of inducing the desired immune response as B cell epitopes. The epitope prediction tools could facilitate the development of vaccines and predict the epitopes (Sitompul et al., 2012). Vaccine production that depends on biochemical experiments can be expensive, time consuming and not always working, although this vaccine formulation of attenuated or inactivated microorganism contains a few hundred of unnecessary proteins for the induction of immunity, that may cause allergenic or reactogenic responses (Lo et al., 2013; Li et al., 2014). Therefore, prediction of epitopes of appropriate protein residues would help in production of peptide vaccine with powerful immunogenic and minimal allergenic effect (Purcell et al., 2007; Reche et al., 2014). Further work is required to clone the gene for these peptides on some carrier vectors in order to collect data on the extent of elicitation of immune response and to ensure that this vaccine will provide long-term protection.

In conclusion, our experiment showed two epitopes (DKAVNVYTSSQT and NMPKDKKEACAKAPEA) proposed to be a peptide vaccine against NDV. Further study is also needed to identify the interaction between these peptides with T cells and antibodies.

ACKNOWLEDGEMENTS

The authors would like to thank the Ministry of Research, Technology and Higher Education of the Republic of Indonesia through Master Degree to Doctorate for Superior Scholarship Program (PMDSU), grant number 2146/D3/PBG/2017.

REFERENCES


Multi-epitope peptide vaccine prediction against Newcastle disease virus using immuno-informatics...


Paper received 12.05.2019; accepted for publication 07.07.2019

**Correspondence:**
Fedik Abdul Rantam,
Jl. Mulyorejo, Kampus C Unair,
Surabaya, 60115, Indonesia,
email: fedik-a-r@fhk.unair.ac.id

8

BJVM, **xx**, No **x**