



# Strategic construction of mRNA vaccine derived from conserved and experimentally validated epitopes of avian influenza type A virus: a reverse vaccinology approach

## Leana Rich Herrera-Ong

Department of Biochemistry and Molecular Biology, College of Medicine, University of the Philippines Manila, Metro Manila, Philippines

Received: January 22, 2023  
 Accepted: March 31, 2023

Corresponding author:

Leana Rich Herrera-Ong, RCh, PhD  
 Department of Biochemistry and Molecular Biology, College of Medicine, University of the Philippines Manila, 670 Padre Faura St., Ermita, Manila, Philippines  
 Tel: +63-2-85264197, +639433124875  
 Fax: +63-2-85264197  
 E-mail: lmherrera2@up.edu.ph; leanaherrera@yahoo.com

No potential conflict of interest relevant to this article was reported.

I would like to thank my family and my co-faculty members at the Department of Biochemistry and Molecular Biology in University of the Philippines Manila for their support and encouragement in completing this task.



© Korean Vaccine Society.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<https://creativecommons.org/licenses/by-nc/4.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

**Purpose:** The development of vaccines that confer protection against multiple avian influenza A (AIA) virus strains is necessary to prevent the emergence of highly infectious strains that may result in more severe outbreaks. Thus, this study applied reverse vaccinology approach in strategically constructing messenger RNA (mRNA) vaccine construct against avian influenza A (mVAIA) to induce cross-protection while targeting diverse AIA virulence factors.

**Materials and Methods:** Immunoinformatics tools and databases were utilized to identify conserved experimentally validated AIA epitopes. CD8<sup>+</sup> epitopes were docked with dominant chicken major histocompatibility complexes (MHCs) to evaluate complex formation. Conserved epitopes were adjoined in the optimized mVAIA sequence for efficient expression in *Gallus gallus*. Signal sequence for targeted secretory expression was included. Physicochemical properties, antigenicity, toxicity, and potential cross-reactivity were assessed. The tertiary structure of its protein sequence was modeled and validated *in silico* to investigate the accessibility of adjoined B-cell epitope. Potential immune responses were also simulated in C-ImmSim.

**Results:** Eighteen experimentally validated epitopes were found conserved (Shannon index <2.0) in the study. These include one B-cell (SLLTEVETPIRNEWGCR) and 17 CD8<sup>+</sup> epitopes, adjoined in a single mRNA construct. The CD8<sup>+</sup> epitopes docked favorably with MHC peptide-binding groove, which were further supported by the acceptable  $\Delta G_{\text{bind}}$  (-28.45 to -40.59 kJ/mol) and Kd (<1.00) values. The incorporated Sec/SPI (secretory/signal peptidase I) cleavage site was also recognized with a high probability (0.964814). Adjoined B-cell epitope was found within the disordered and accessible regions of the vaccine. Immune simulation results projected cytokine production, lymphocyte activation, and memory cell generation after the 1st dose of mVAIA.

**Conclusion:** Results suggest that mVAIA possesses stability, safety, and immunogenicity. *In vitro* and *in vivo* confirmation in subsequent studies are anticipated.

**Keywords:** Avian influenza A virus, Vaccinology, Messenger RNA, Chickens, Epitopes

## Introduction

Avian influenza, commonly known as bird flu, is caused by influenza A viruses (IAV) that circulate in humans, horses, dogs, swine, and birds. Wild aquatic birds, shore-birds, and wild waterfowl are considered reservoir hosts of IAV [1]. Avian influenza has

captured the attention of international and local communities over the years due to its devastating consequences to the health of wild birds, poultry, and livelihood of farmers. It has become a public health concern because of sporadically identified human cases. A major concern is the emergence of highly infectious strains that may result in pandemics.

Strains of avian influenza viruses can be generally classified into two categories, depending on the disease severity in poultry. The low pathogenic avian influenza (LPAI) viruses cause either no signs of disease to mild disease which can be observed as ruffled feathers and marked decrease in egg production. Some LPAI viruses can mutate into highly pathogenic avian influenza (HPAI) viruses which can result in severe disease that affects multiple organs. Mortality is up to 100% in infected chickens, often within few days of incubation [2]. IAV is more commonly divided into subtypes based on the type of two proteins expressed on its surface: hemagglutinin (HA) and neuraminidase (NA). In humans, there are 18 subtypes of HA and 11 subtypes of NA, while 16 HA and nine NA subtypes have been identified in birds. There are many possible combinations of HA and NA that result in subtype of IAV. An A(H5N1) strain has HA 5 and NA 1 type of proteins on its surface. Currently, all known subtypes of IAV can infect birds, except two subtypes that have only been found in bats: A(H18N11) and A(H17N10). While most of the circulating subtypes A(H5) and A(H7) viruses are LPAI, and only few have been classified as HPAI, both can spread rapidly through poultry flocks [3].

IAV is the only species of the genus *Alphainfluenzavirus*, family *Orthomyxoviridae*. Its genome is comprised of eight negative-sense, single-stranded RNA segments which encode for 10 viral proteins namely: polymerase basic 2 (PB2), polymerase basic 1 (PB1), polymerase acidic protein (PAP), HA, nucleoprotein (NP), NA, matrix-1 protein (M1), matrix-2 protein (M2), non-structural protein 1, and non-structural protein 2. Seven new proteins encoded by the viral genome have been identified in the past several years. Each RNA segment is wrapped around NP monomers that form a ribonucleoprotein complex with PAP, PB1, and PB2 for viral transcription and replication. The complexes are enclosed by a lipid envelope that is layered by M1 proteins internally, and embedded by HA, NA, and M2 proteins. HA protein facilitates viral entry by binding to the sialic acid receptor of the host cell. NA protein promotes the release of viral progeny from the host cell [4].

So far, the largest HPAI epidemic season was observed

from 2021–2022 in Europe. This involved 2,733 HPAI infections of wild birds, 2,398 outbreaks in poultry, and 168 detections in captive birds which occurred in 36 European countries. A total of 46 million birds were culled in the affected establishments [5]. The new H5N1 HPAI is responsible for the current outbreak since early 2022; and it has effectively displaced other H5 HPAI viruses. In fact, the ongoing HPAI H5N1 outbreak has led to the death of over 37 million poultry in the United States alone [6]. In the Philippines, new outbreaks have been recorded from different provinces across Luzon, including Pampanga, Bulacan and Laguna [7].

In countries with good veterinary infrastructure, mass HPAI poultry vaccination has been successfully accomplished. Currently available platforms of avian influenza vaccines include attenuated wild-type avian influenza A viruses (AIAV), protein subunit HA, adjuvanted HA DNA vaccine, and recombinant live virus vectors expressing HA and NA gene inserts. However, aside from the challenges with cost production and limited cold chain capacity, these currently available vaccines are not commonly used because of their reduced vaccination efficacy over time [8]. Unequivocally, there is a need for constant reformulation of avian influenza vaccines due to antigenic drift and antigenic shift of AIAV in the field. Therefore, development of vaccines that can cover multiple antigens across multiple currently circulating strains can be a more practical strategy. Moreover, avian influenza vaccines targeting multiple bird species would be of greater value, most especially in South East Asia wherein some species like ducks are often carrying the virus without expressing any clinical symptoms. Evidently, the ongoing diversification of IAV challenges the conventional method of vaccine formulation which is time-consuming, production-limited, and strain-specific. Keeping in mind that cost is another big hurdle to be considered when it comes to individual bird vaccination, vaccines that could confer at least 6-month protection after a single dose could significantly minimize cost required for mass vaccination. Thus, better strategic approaches and more applicable vaccine platforms against intracellular pathogens such as AIAV should be explored.

This study aims to construct a messenger RNA (mRNA) vaccine which incorporates conserved and experimentally validated epitopes from multiple antigens of AIAV. To induce cross-protection against various AIAV strains, conserved regions can be strategically identified in AIAV protein sequences from public databases. The use of highly conserved sequences is an important factor for vaccine development in

order to circumvent immune epitope evasion in rapidly mutating viruses such as AIAV. In addition, targeting multiple virulence factors may induce broader and more potent immune responses against AIAV. This study will also explore the essential elements required in constructing mRNA as a vaccine platform to incorporate the putative advantages that therapeutic mRNA molecules possess. These include the forestalling of host genome integration risk, lower toxicity, and the ability to get host cell to produce viral proteins for easier, faster, and cheaper manufacturing processes than the traditional vaccines [9].

The development of immunoprophylactic and immunotherapeutic agents has become more cost-effective and time-efficient in the immunoinformatics era. The advent of coronavirus disease 2019 (COVID-19) pandemic has demonstrated the cutting-edge and paramount importance of immunoinformatics when biotechnology and pharmaceutical companies around the world have successfully and rapidly developed vaccines against severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). With the application of appropriate databases and immunoinformatics tools, coupled by a strategic approach, conserved and experimentally validated epitopes in AIAV antigens can be identified, and an mRNA construct can be designed to comprise essential elements that a potent AIAV vaccine must possess. Results from this study can serve as premises in formulating effective mRNA vaccine that can potentially induce longer-lasting immunity against multiple AIAV strains.

## Materials and Methods

### Retrieval of experimentally validated epitopes and the identification of conserved regions in AIA protein sequences

All experimentally validated avian influenza A (AIA) T-cell and B-cell epitopes with positive assay results were retrieved from the Immune Epitope Database and Analysis Resource (IEDB) on November 4, 2022 (<http://tools.iedb.org/conservancy/>). The search criteria were limited to linear epitopes of IAV (ID:11320) with birds (ID:8782) as host. Manually curated protein sequences corresponding to epitopes obtained from IEDB were generated from the UniProt database (<https://www.uniprot.org/>). Inclusive lengths of HA, matrix protein 1 (MP1), matrix protein 2 (MP2), non-structural protein 1 (NSP1), NP, PAP, and RNA-directed RNA polymerase catalytic subunit (RRPCS) are 564-568, 252, 97, 230-237, 498, 716, and 757, respectively. Retrieved protein sequences were aligned in

Clustal Omega server (<https://www.ebi.ac.uk/Tools/msa/clustalo/>) which were used as input in the Protein Variability Server tool (PVS; <http://imed.med.ucm.es/PVS/>). Fragments with Shannon index <2.0, and length greater than or equal to 9 residues were identified as conserved. All variable residues were masked in the resulting sequence. To determine which among the experimentally validated epitopes are conserved, epitope that has 100% residue identity overlap with any of the conserved fragments of their respective AIAV antigen were identified using the Epitope Conservancy Analysis tool (<http://tools.iedb.org/conservancy/>). Conserved epitopes were kept for subsequent analysis.

Currently, with the settings employed in this study, there is only one (1) experimentally validated AIA T-helper (Th) cell epitope available in IEDB. Unfortunately, this epitope was not found to be conserved upon alignment with the conserved sequences of AIAV antigens identified in this study. To circumvent the scarcity of Th epitopes, two universal Th cell epitopes were included in the vaccine construct. Th epitope TET (QYIKANSKFIGIT) was derived from the sequence of tetanus neurotoxin of *Clostridium tetani*. It was recently demonstrated that the incorporation of TET in a vaccine was able to induce humoral and cellular immunity in chickens and mice [10]; and was also utilized as marker to identify vaccinated from unvaccinated chickens [11]. Thus, induced immunity through this epitope may also enable the distinction of vaccinated chickens from infected ones. Another universal epitope (AQFVRALSMQAAE) known to activate Th cells was derived from pan-allelic DR-binding sequence (PADRE). Studies have demonstrated its ability to activate Th cells regardless of chicken haplotypes [12].

### Assessment of potential toxicity and cross-reactivity of conserved epitopes

The potential toxicity of each epitope was evaluated in ToxinPred server (<https://webs.iitd.edu.in/raghava/toxinpred/index.html>). Epitopes with negative support vector machine (SVM) scores were classified as non-toxin. To avoid potential cross-reactivity induced by epitopes against the proteome of the target host, sequence similarity search for each epitope was conducted through Protein BLAST against non-redundant protein sequences of birds (taxid:8782) in the National Center for Biotechnology Information protein databases (<https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins>). This algorithm automatically adjusts for shorter input sequence. All model sequences, non-redundant RefSeq pro-

teins, and uncultured/environmental sample sequences were excluded. Each conserved epitope was assessed for exact hits. Lengths of peptides presented by major histocompatibility complex (MHC) I can range from 8 to 12 residues; thus, epitopes with 100% sequence identity match for  $\geq 8$  residues in any of the avian protein sequences in the database were excluded to prevent plausible provocation of immune response against the host. Due to short fragment length inputs, resulting E-values for each epitope were too large (3.3–400); therefore, this parameter was not included as a criterion in identifying hits.

### Docking conserved epitopes with dominant MHC I and MHC II alleles of chickens

Chickens express one dominant classical MHC I and MHC II, known as BF2 haplotype [13] and BL2\*02, respectively [14]. The PDB crystal structures of MHC I BF2\*2101 (PDB ID:3BEV) and MHC II BL2\*02 (PDB ID:7APZ) were downloaded from RSCB Data Bank (<https://www.rcsb.org/>). Unnecessary structures were removed from the PDB files. All conserved cytotoxic T-cell (Tc) epitopes identified in the study were docked with BF2\*2101, while the other two universal Th cell epitopes were docked with BL2\*02 in GalaxyPepDock server (<https://galaxy.seoklab.org/cgi-bin/submit.cgi?type=PEPDOCK>). Resulting docked models with the highest TM scores were retrieved and further refined using the RefineComplex tool in GalaxyWEB server (<https://galaxy.seoklab.org/cgi-bin/submit.cgi?type=COMPLEX>). The favorability of binding between an epitope and its corresponding MHC binder was further investigated by calculating the binding free energy ( $\Delta G_{\text{bind}}$ ) and the dissociation constant (Kd) of complex formation in the PRODIGY web-server (<https://wenmr.science.uu.nl/prodigy/>). Temperature was set at 43°C as this is reported to be the highest normal body temperature of birds [15].

### In silico construction of AIA mRNA vaccine and the optimization of its ORF nucleic acid sequence

The essential sequence elements required for efficient translation of mRNA vaccine in *Gallus gallus* (used as representative host) were included in the mRNA vaccine construct against avian influenza A (mVAIA). Generally, mVAIA was designed to contain a 5' cap, a 5' untranslated region (UTR), an open reading frame (ORF), a 3' UTR, and a poly(A) tail sequence. This study suggests the use of 5' cap 1 (m<sup>7</sup>GpppN1m) in the mRNA construct to increase translational efficiency. The 5' UTR sequence was derived from the nucleic acid sequence of

*Gallus gallus* troponin C (NM\_205133.2) which includes a Kozak consensus sequence and a start codon. To prevent the formation of secondary structures at the 5' UTR end of the vaccine, a poly(A) sequence of 12 nucleotides was inserted as part of the leader sequence after the 5' cap 1 [16]. Moreover, this study recommends the use of modified nucleosides in synthesizing mRNA to reduce host immunostimulatory activities against the vaccine. Studies have shown that synthetic mRNAs containing 5-methylcytidine (m5C), N6-methyladenosine (m6A), 5-methyluridine (m5U), and 2-thiouridine (s2U) can diminish the stimulation of inherent immune reactions [17]. The whole ORF sequence of mVAIA comprises the tissue-type plasminogen activator (PLAT) signal sequence of *Gallus gallus* (Uniprot ID:E1C209) which is conjugated to the series of the experimentally validated AIA epitopes. The epitopes were connected via AAY and GPGPG linkers starting from the linear B-cell epitope from the N-terminus to the series of CD8<sup>+</sup> epitopes at the C-terminus. The two universal Th epitopes were positioned between the linear B-cell epitope and the series of CD8<sup>+</sup> epitopes. The 3' UTR sequence of the vaccine was derived from the nucleotide sequence of *Gallus gallus* troponin C (NM\_205133.2) which includes a stop codon (UAA). Lastly, the poly(A) tail sequence of mVAIA was designed to contain 100nts which has been shown to be ideal for therapeutic mRNA [18]. To incorporate the ORF sequence in the mRNA construct, the protein sequence of the ORF was back-translated using EMBOSS Backtranseq tool in EMBL-EBI ([https://www.ebi.ac.uk/Tools/st/emboss\\_backtranseq/](https://www.ebi.ac.uk/Tools/st/emboss_backtranseq/)). *Gallus gallus* was selected as host in the codon usage table parameter. The translation efficiency of the ORF was expressed as Codon Adaptation Index (CAI). The CAI and the percentage GC content of optimized ORF sequence for expression in *Gallus gallus* was calculated in VectorBuilder Codon Optimization tool (<https://en.vectorbuilder.com/tool/codon-optimization.html>).

### Evaluation of the secondary structure of nucleotide sequence and in silico construction of the DNA plasmid encoding mVAIA

The mFold web server (<http://www.unafold.org/>) was used to calculate and predict the minimum free energy (MFE) for the optimal secondary structure of nucleotide sequence of mVAIA. GU pairs and isolated bases were avoided at the end of the helices. *In silico* cloning of mVAIA was conducted using the GeneSmart Design tool (<https://www.genscript.com/gene-and-plasmid-construct-design.html#video2>). Restriction

tion enzymes StyI and BgIII were selected as insertion sites for the whole gene encoding the DNA sequence of mVAIA. Restriction sites for the aforementioned endonucleases were also avoided during the sequence optimization of mVAIA.

### Targeting mVAIA for secretion and the assessment of its physicochemical properties

To increase the chances of the protein product being secreted, mVAIA was also designed to contain a signal peptidase sequence for its translocation to the endoplasmic reticulum. The SignalIP 6.0 server (<https://services.healthtech.dtu.dk/service.php?SignalP>) was utilized to identify the potential signal peptidase cleavage site in the full protein sequence. SignalIP predicts the presence of secretory/signal peptidase I (Sec/SPI) in a sequence which are standard secretory signal peptides that are recognized and transported by the Sec translocon, and subsequently cleaved by SPI. Terminal peptide sequence predicted with a very high probability of being cleaved off will be less likely included in the secreted protein sequence form of the vaccine.

The secreted protein sequence of mVAIA was further evaluated *in silico* for its antigenicity, allergenicity, cross-reactivity, toxicity, and physicochemical properties. The potential antigenicity of the secreted form was determined in Vaxijen ver. 2.0 (<http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html>) antigenicity tool by setting the threshold to >0.4 for viral antigens. Vaxijen's prediction accuracy was estimated to be 83% for cross-validation, and 80% for external test set. Potential allergenicity was also evaluated in AllerTop ver. 2.0 (<https://www.ddg-pharmfac.net/AllerTOP/feedback.py>). AllerTOP outperforms other allergen-prediction servers with its 94% sensitivity rate [19]. In order to prevent potential cross-reactivity of the vaccine with the proteome of target host (*Gallus gallus*), the whole protein sequence of mVAIA was evaluated for significant sequence similarity hits with the non-redundant protein sequence database in Protein BLAST (<https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins>). Sequences including models, non-redundant RefSeq, and uncultured/environmental samples were excluded. The potential toxicity of the vaccine's secreted protein sequence form was assessed in ToxinPred server; wherein stretches of decamers with negative SVM scores were classified as non-toxin. As crucial as the previous steps, the physicochemical properties of the translated sequence were estimated in ExPASy ProtParam tool (<https://web.expasy.org/protparam/>) to facilitate future isolation, purification, storage procedure, and

characterization of mVAIA as a vaccine candidate.

### Secondary structure composition and accessibility of the adjoined B-cell epitope

The conserved linear B-cell epitope (SLLTEVETPIRNEW-GCR) adjoined in mVAIA should be preferably in a disordered and paratope-accessible region of the vaccine antigen, for more effective B-cell receptor (BCR) binding. Thus, this study determined the secondary structure composition of the secreted protein sequence of mVAIA in order to identify regions of random coils in GOR4 web tool ([https://npsa-prabi.ibcp.fr/cgi-bin/npsa\\_automat.pl?page=npsa\\_gor4.html](https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_gor4.html)). The tertiary structure model of the protein was predicted using GalaxyTBM tool (<https://galaxy.seoklab.org/cgi-bin/submit.cgi?type=TBM>) to assess the accessibility of adjoined linear B-cell epitope. The model structure was further refined using GalaxyRefine2 tool (<https://galaxy.seoklab.org/cgi-bin/submit.cgi?type=REFINE2>), and validated *in silico* using a myriad of three-dimensional (3D)-structural-model-validation tools. The percentage of residues lying within the favored and the disallowed regions was determined using Zlab Ramachandran Plot sever (<https://zlab.umassmed.edu/bu/rama/index.pl>). The quality factor for the non-bonded atomic interactions in the model structure, calculated as ERRAT score, was determined in UCLA (University of California, Los Angeles) SAVES ver. 6.0 server (<https://saves.mbi.ucla.edu/>). The structure model was viewed using iCn3D structure viewer (<https://www.ncbi.nlm.nih.gov/Structure/icn3d/>). Ellipro structure-based B-cell epitope prediction tool in IEDB was utilized to investigate the accessibility and flexibility of adjoined linear B-cell epitope in the tertiary structure of the secreted protein sequence of mVAIA. Ellipro predicts epitopes based on flexibility and solvent-accessibility of residues in a protein sequence [20].

### *In silico* immune simulation of the secreted protein sequence of mVAIA

The potential immunogenicity of mVAIA was further investigated *in silico* using the secreted protein sequence of the vaccine (excluding the secretory signal peptide sequence) in C-ImmSim server (<https://kraken.iac.rm.cnr.it/C-IMMSIM/index.php?page=1>). This tool simulates immune responses against an antigenic amino acid sequence and was utilized in modeling immune responses from Epstein-Barr virus [21] and human immunodeficiency virus infection [22]. In this study, simulation parameters include 3 times injection of

vaccine with 28 days interval at time steps set 1, 84, and 168, and 200 simulation steps to accommodate time steps.

**Ethics approval**

This study did not conduct methods which include human participants, human samples, animal models, or any animal test system. Thus, ethics review was not deemed necessary.

**Results**

**Experimentally validated epitopes and the conserved regions in AIA protein sequences**

The search criteria used for the retrieval of experimentally validated epitopes from IEDB resulted in a total of 43 epitopes which are comprised of 22 CD8<sup>+</sup>, 1 CD4<sup>+</sup>, and 20 linear B-cell epitopes. A total of 769 manually curated AIA protein sequences, including HA (103), MP1 (113), MP2 (115), NSP1 (97), NP (159), PAP (91), and RRPCS (91) were retrieved from Swissprot KB. Analysis of aligned sequences in PVS resulted in conserved fragments of more than or equal to nine residues with Shannon variability index <2.0. The methods and parameters used in this study were able to identify two linear B-cell epitopes and 21 CD8<sup>+</sup> epitopes that have 100% sequence identity match with the conserved fragments of different AIA antigens. Upon assessment, the toxicity scores of conserved epitopes

**Table 1.** Conserved experimentally validated epitopes of avian influenza A

| Assay classification           | Influenza A antigen | Peptide sequence          |           |
|--------------------------------|---------------------|---------------------------|-----------|
| B-cell                         | Matrix protein 2    | SLLTEVETPIRNEWGCR         |           |
| T-cell class I MHC restriction | Hemagglutinin       | TIGCEPKYV                 |           |
|                                |                     | KTRPILSPL                 |           |
|                                | Matrix protein 1    | ILGFVFTL                  |           |
|                                |                     | MRTIGTHP                  |           |
|                                |                     | VETYVLSI                  |           |
|                                |                     | Nucleoprotein             | KRGINDRNF |
|                                |                     |                           | PKKTGGPIY |
|                                |                     |                           | AEIEDLIFL |
|                                |                     |                           | VMELIRMI  |
|                                |                     |                           | EDLRVSSFI |
|                                |                     |                           | GRRTRIAIY |
|                                |                     |                           | NATEIRASV |
|                                |                     |                           | PTFSVQRNL |
|                                |                     |                           | VGTMVMELI |
|                                |                     |                           | YDKEEIRRI |
|                                |                     | AVKGVGTMVME               |           |
|                                |                     | Polymerase acidic protein | GRRKTNLY  |

MHC, major histocompatibility complex.

range from -1.23 to -0.34. All epitopes were classified as non-toxin. The study found that one of the validated epitopes, linear B-cell epitope WNDNT in NSP1, had 100% sequence identity and 100% sequence coverage hit with an avian protein sequence in the database. Thus, pentapeptide WNDNT was excluded from the final set of conserved epitopes to be included in the vaccine. All overlapping epitopes of an antigen were merged with other longer epitopes as the representative fragment (Table 1).

**Epitopes docked with dominant chicken MHC I and MHC II molecules**

The conserved CD8<sup>+</sup> epitopes were docked with the structure of dominant chicken MHC I while the TET and PADRE epitopes were docked with the dominant MHC II molecule of *Gallus gallus*. All docking procedures were conducted in Galaxy-PepDock which resulted in epitope-MHC complex formation having acceptable  $\Delta G_{bind}$  and Kd values (Table 2). The estimated accuracy of docked structures ranges from 0.727 to 0.939. Fig. 1 shows that each epitope is docked on the binding groove of the respective dominant chicken MHC type utilized in the study.

**Table 2.** Docking accuracy, binding energy, and dissociation constants for epitope-MHC complex formation

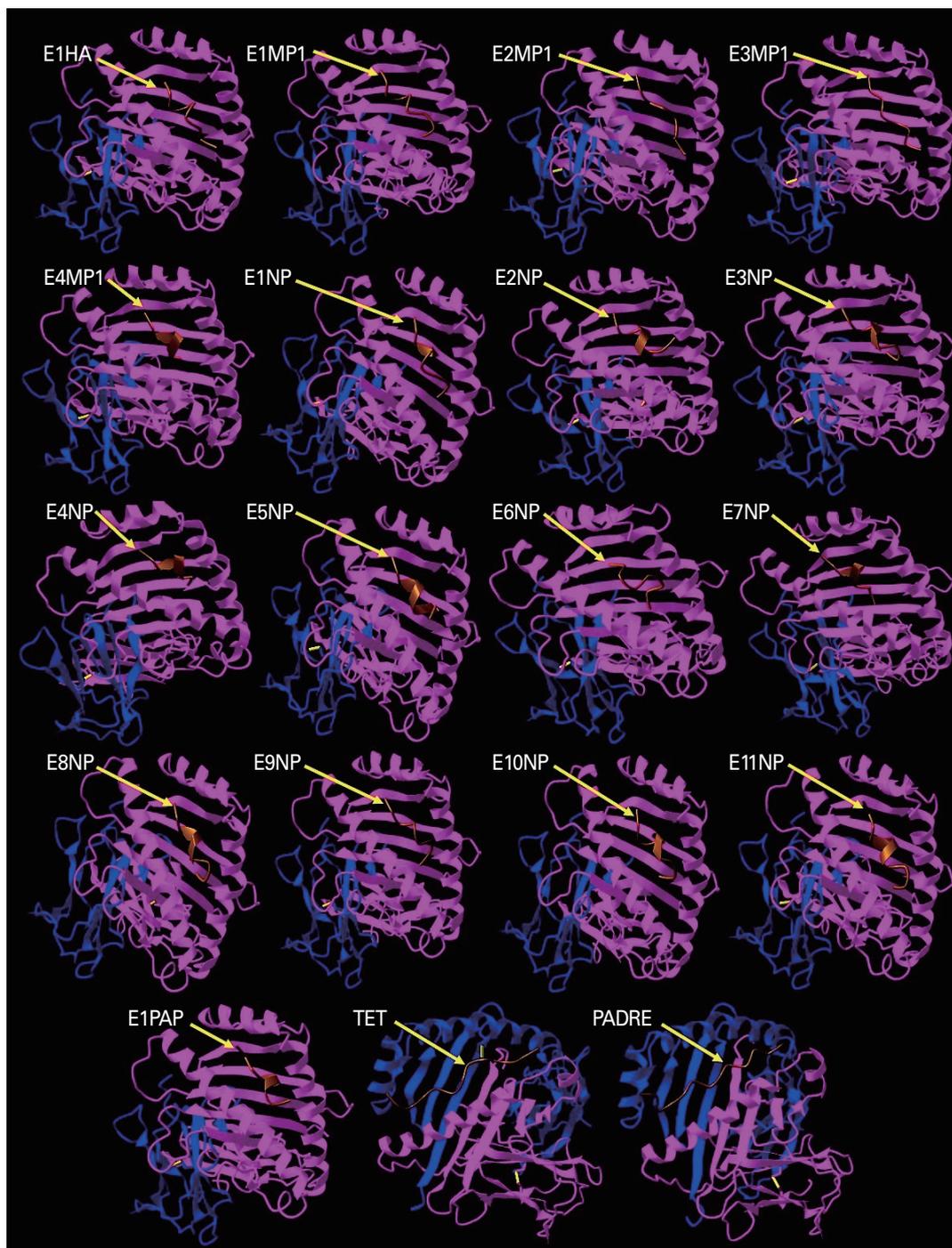
| MHC type         | Sequence      | Epitope code | Estimated docking accuracy | $\Delta G_{bind}$ (kJ/mol) | Kd (mol/dm <sup>3</sup> ) |
|------------------|---------------|--------------|----------------------------|----------------------------|---------------------------|
| Chicken BF2 2101 | TIGCEPKYV     | E1HA         | 0.868                      | -36.82                     | 7.90E-07                  |
|                  | KTRPILSPL     | E1MP1        | 0.912                      | -34.73                     | 1.70E-06                  |
|                  | ILGFVFTL      | E2MP1        | 0.911                      | -32.23                     | 4.90E-06                  |
|                  | MRTIGTHP      | E3MP1        | 0.727                      | -37.24                     | 6.90E-07                  |
|                  | VETYVLSI      | E4MP1        | 0.899                      | -35.56                     | 1.40E-06                  |
|                  | KRGINDRNF     | E1NP         | 0.843                      | -38.91                     | 3.40E-07                  |
|                  | PKKTGGPIY     | E2NP         | 0.939                      | -33.89                     | 2.50E-06                  |
|                  | AEIEDLIFL     | E3NP         | 0.939                      | -33.89                     | 2.50E-06                  |
|                  | VMELIRMI      | E4NP         | 0.882                      | -30.54                     | 8.40E-06                  |
|                  | EDLRVSSFI     | E5NP         | 0.906                      | -33.05                     | 3.70E-06                  |
|                  | GRRTRIAIY     | E6NP         | 0.768                      | -31.38                     | 6.60E-06                  |
|                  | NATEIRASV     | E7NP         | 0.853                      | -40.59                     | 1.80E-07                  |
|                  | PTFSVQRNL     | E8NP         | 0.917                      | -38.07                     | 5.40E-07                  |
|                  | VGTMVMELI     | E9NP         | 0.91                       | -28.45                     | 2.00E-05                  |
|                  | YDKEEIRRI     | E10NP        | 0.883                      | -33.05                     | 3.40E-06                  |
|                  | AVKGVGTMVME   | E11NP        | 0.921                      | -40.59                     | 1.90E-07                  |
| GRRKTNLY         | E1PAP         | 0.866        | -39.33                     | 3.00E-07                   |                           |
| Chicken BL2 02   | QYIKANSKFIGIT | TET          | 0.84                       | -48.53                     | 9.00E-09                  |
|                  | AQFVRALSMQAAE | PADRE        | 0.82                       | -46.86                     | 1.70E-08                  |

MHC, major histocompatibility complex.

**Construction of avian influenza A mRNA vaccine *in silico***

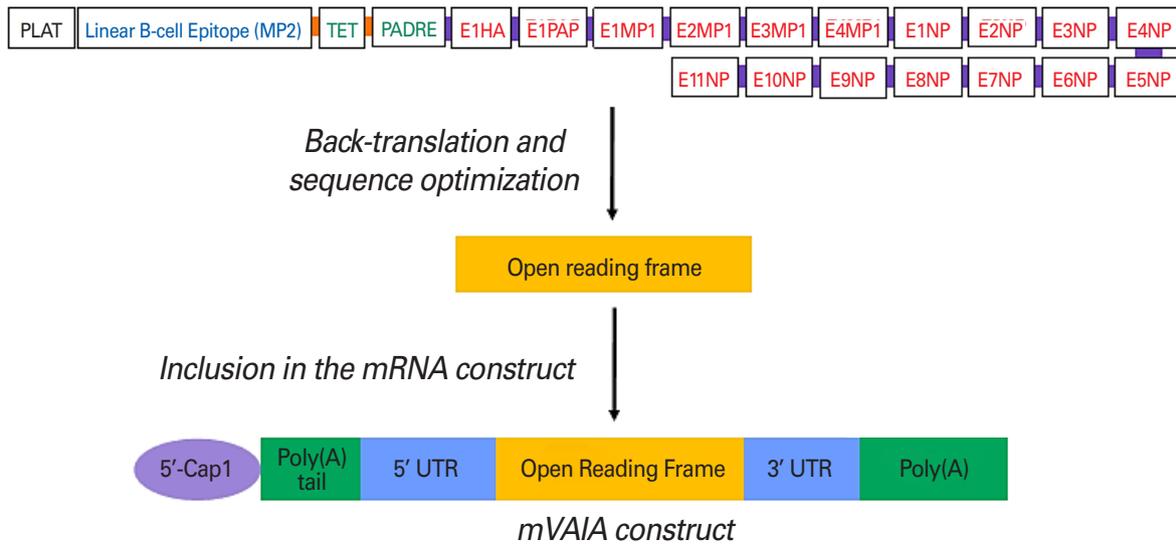
Fig. 2 depicts the schematic representation of the mVAIA. Epitopes were arranged starting from the linear B-cell epitope of MP2 (text in blue), followed by the two universal CD4<sup>+</sup> epitopes for chickens (texts in green), and the series of CD8<sup>+</sup> epitopes of HA, PAP, MP1, and NP (texts in red). All overlap-

ping CD8<sup>+</sup> epitopes were merged with the longer epitopes per antigen. The linear B-cell epitope and two universal Th epitopes are connected by GPGPG linkers (orange bars), while CD8<sup>+</sup> epitopes are adjoined via AAY linkers (purple bars). There was only one AIA CD4<sup>+</sup> epitope available in IEDB at the time of data collection; however, the epitope was

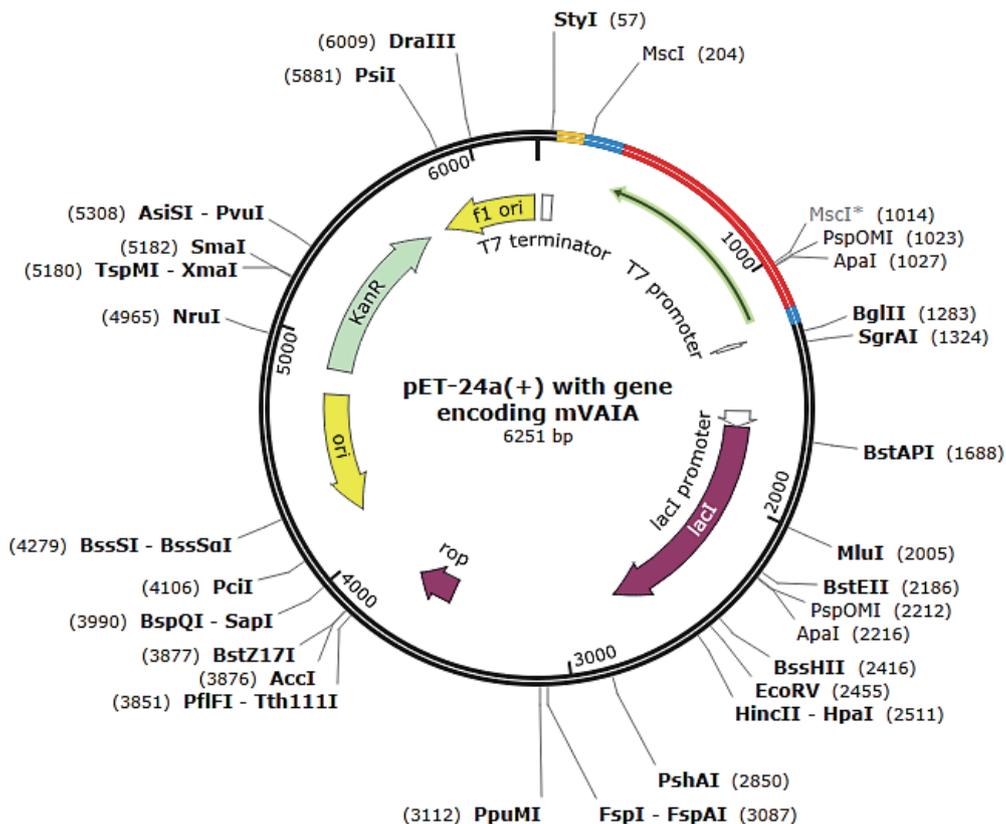


**Fig. 1.** Conserved avian influenza A epitopes docked on chicken major histocompatibility complex (MHC) binding groove. Tertiary models of MHC molecule protein subunits are depicted as blue and magenta ribbon structures. All epitopes docking exactly on the peptide-binding grooves are depicted as brown ribbon structures where yellow arrows are pointed to.

Conserved and experimentally validated epitopes adjoined with PLAT, TET, and PADRE



**Fig. 2.** Schematic representation for the construction of messenger RNA (mRNA) vaccine construct against avian influenza A (mVAIA). The peptides are aligned starting from the plasminogen activator (PLAT) sequence in black text, linear B-cell epitope of matrix protein 2 (MP2) in blue text, TET and PADRE universal CD4<sup>+</sup> epitopes in green text; and the series of CD8<sup>+</sup> epitopes of hemagglutinin (HA), polymerase acidic protein (PAP), and matrix protein 1 (MP1) antigens in red texts which are all adjoined by AAY and GPGPG linkers depicted as orange and purple bars in between boxes, respectively. The adjoined amino acid sequence was back-translated to its corresponding nucleotide sequence, represented by the yellow-gold box, which was inserted to the whole mRNA construct sequence, herein called mVAIA construct.



**Fig. 3.** Schematic representation of the DNA plasmid. The pET-24a(+) contains all the essential elements for *in vitro* transcription of gene that encodes for messenger RNA vaccine construct against avian influenza A (mVAIA).

not found to be conserved in this study. Thus, two universal Th epitopes were added in the vaccine. The partial sequence (TET) of tetanus neurotoxin (QYIKANSKFIGIT) from *Clostridium tetani* (sequence ID: WP\_129031034.1) and the PADRE sequence (AQFVRALSMQAAE) were both included in the construct. The signal peptide from PLAT sequence of *Gallus gallus* (E1C209) was adjoined with the N-terminus of the whole sequence assembly which has a total of 289 amino acid residues. The sequence was reverse translated and optimized to its equivalent RNA sequence to represent the ORF. The optimized sequence has 56.4% GC content and a CAI of 0.93 in *Gallus gallus* as host. Restriction enzyme sites StyI and BglIII were also evaded in the optimized sequence.

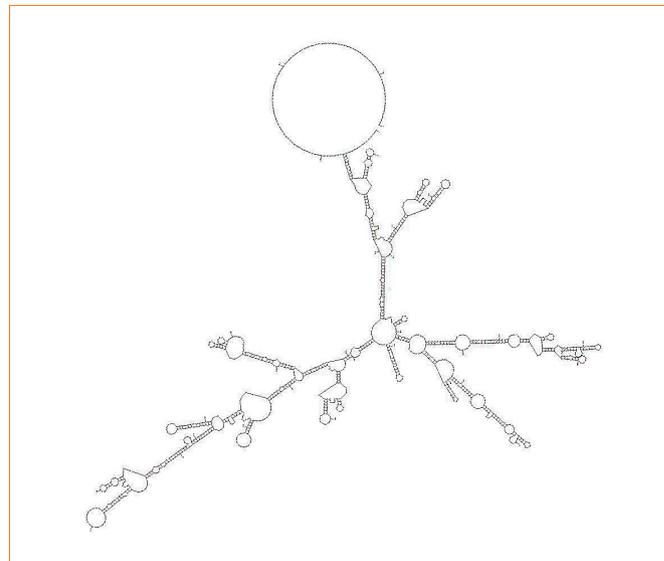
The ORF sequence was included in the mRNA vaccine construct. This study proposes the use of 5' cap 1 (<sup>m7</sup>GpppN1<sup>m</sup>) in synthesizing mVAIA to increase translation efficiency. The UTR of mVAIA were derived from the sequence of *Gallus gallus* troponin C (NM\_205133.2) which is 54 bases long, including the Kozak consensus sequence and a start codon. A poly(A) leader sequence of 12 nucleotides was also included. The 3' UTR sequence is 147 bases long, including the stop codon. Lastly, a poly(A) tail sequence with 100 bases was added at the 3' end of the construct.

### Construction of DNA plasmid encoding mVAIA

Fig. 3 shows pET24a(+) pDNA vector containing a gene encoding for mVAIA (red). Both 5' and 3' UTRs are shown in blue, while the Poly(A) tail is depicted in orange. The green arrow shows the direction of transcription. All essential sequence features for replication of pDNA and transcription of mVAIA gene via T7 RNA polymerase are also indicated. Restriction sites StyI and BglIII, which can be utilized as insertion sites, are located closest to the terminal of the gene encoding mVAIA. To ensure that the correct amino acid sequence is encoded by the nucleotide sequence inserted in the plasmid, the DNA sequence of the gene in the plasmid was translated to its corresponding amino acid sequence. The resulting sequence was evaluated against the original amino acid sequence of mVAIA in the IEDB Conservancy Tool. Results showed 100% sequence similarity between the two amino acid sequences (289/289).

### Thermal stability of the secondary structure of mVAIA

The complete mRNA sequence of the candidate vaccine was evaluated in the mFold web server. The MFE or  $\Delta G_{\text{fold}}$  calculated for its optimal secondary structure (Fig. 4) is -1635.94



**Fig. 4.** The optimal secondary structure of messenger RNA vaccine construct against avian influenza A (mVAIA). The optimized secondary structure contains external loop with single stranded bases at the terminal ends of the sequence. Inner structures contain interior loops, multi-loops, and bulge loops.

kJ/mol which suggests a great thermal stability [23]. Noteworthy, Fig. 4 shows that both 5' and 3' ends are part of an external loop that has single stranded bases which may allow more efficient binding of the translation initiation factors [24].

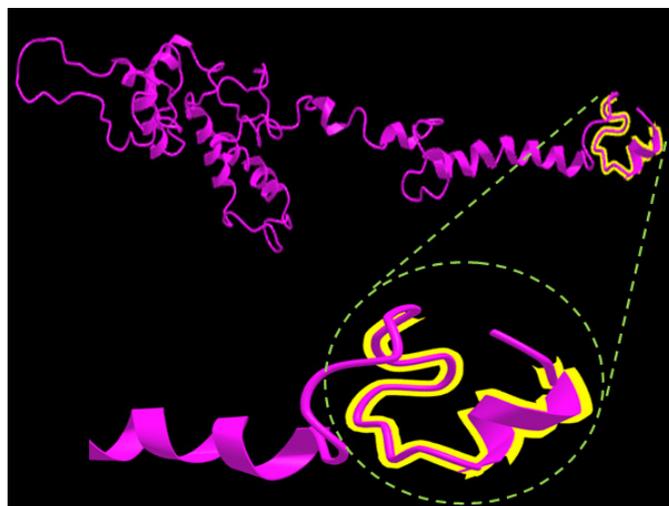
### Cleavage of signal peptide from the translated sequence of mVAIA

SignalP ver. 6.0 server has recognized the adjoined signal peptide sequence Sec/SPI with very high probability (0.964814). The identified cleavage site is located between the 35th and 36th residues of the complete amino acid sequence of mVAIA (289 residues). Fig. 5 depicts the positions of Sec/SPI n (red), h (orange), c (yellow), and the exact position of the cleavage site (broken line in green) in the sequence. The epitopes encoded by the ORF sequence were further rearranged to explore and analyze which adjacent sequence with respect to the PLAT sequence would result to higher probability of being recognized. Analysis of rearranged sequences showed that when PLAT is adjacent to AAY linker, the probability of identifying Sec/SPI sequences decreases to 0.8171. In another rearrangement, the probability drops down to 0.7311 when PLAT is adjacent to CD8<sup>+</sup> epitope (TIGECCPKYV). Thus, the initial arrangement was followed wherein the linear B-cell epitope is adjacent to the PLAT sequence. Residues 1–35 were removed from the complete



Fig. 7 shows the predicted tertiary structure model of the vaccine, highlighting the position of the linear B-cell epitope adjoined in the vaccine (highlighted in yellow). Analysis of Ramachandran plot suggests that the predicted tertiary structure model is valid (Fig. 8). The Ramachandran plot of 3D structures was generated from the amino acid torsion angles of the protein's tertiary structure model. Residues depicted as green spots are within the highly favored region (94.841%), residues depicted as brown spots are within favored region (4.762%), and only one residue is within the disallowed region (0.397%). For the best quality model, 90% of the torsion angles of a tertiary structure should fall in the favored and allowed regions [26]. Moreover, the predicted model has an ERRAT score of 77.88 which indicates a good-quality tertiary structure model. ERRAT calculates the percentage of the protein for which the error value falls below the 95% rejection limit. A model structure with ERRAT score >50% is considered to be of a high-quality [27].

The tertiary structure model of the secreted protein se-

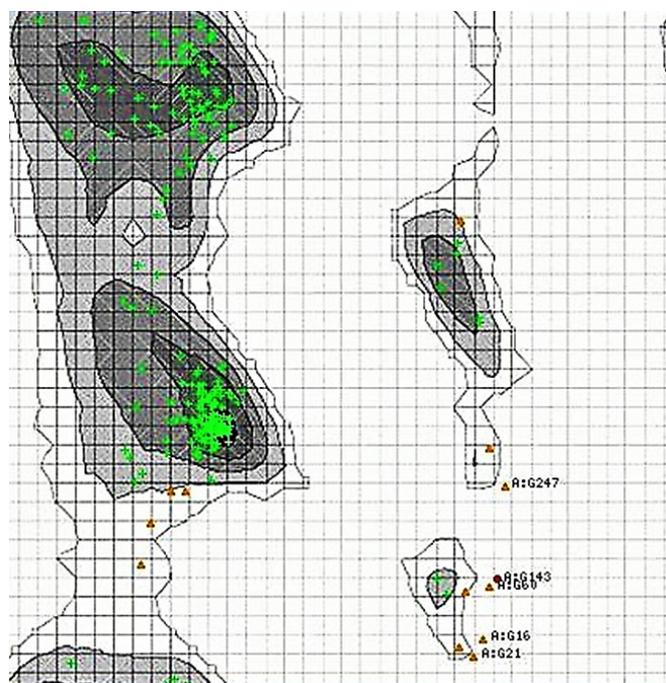


**Fig. 7.** Tertiary structure model of messenger RNA vaccine construct against avian influenza A (mVAIA). The three-dimensional model of the secreted protein structure of mVAIA is shown in the upper portion. The adjoined linear B-cell epitope SLLTEVETPIRNEWGCR was highlighted in yellow, and was further enlarged as depicted in the lower portion of the figure.

quence of mVAIA was employed in a B-cell epitope prediction tool to assess whether the sequence of the adjoined linear B-cell epitope can still be recognized as a B-cell epitope though it is linked to another sequence, forming a tertiary structure. Results from the analysis in Ellipro showed that the residues of the adjoined linear B-cell epitope (SLLTEVETPIRNEWGCR) are all accessible as linear and discontinuous epitope in the tertiary structure of the vaccine with protrusion scores 0.8 and 0.821, respectively (Table 3). It has been regarded that higher protrusion index indicates greater solvent accessibility of the residues [20].

**Immune simulation**

The secreted protein sequence of mVAIA was used as antigen

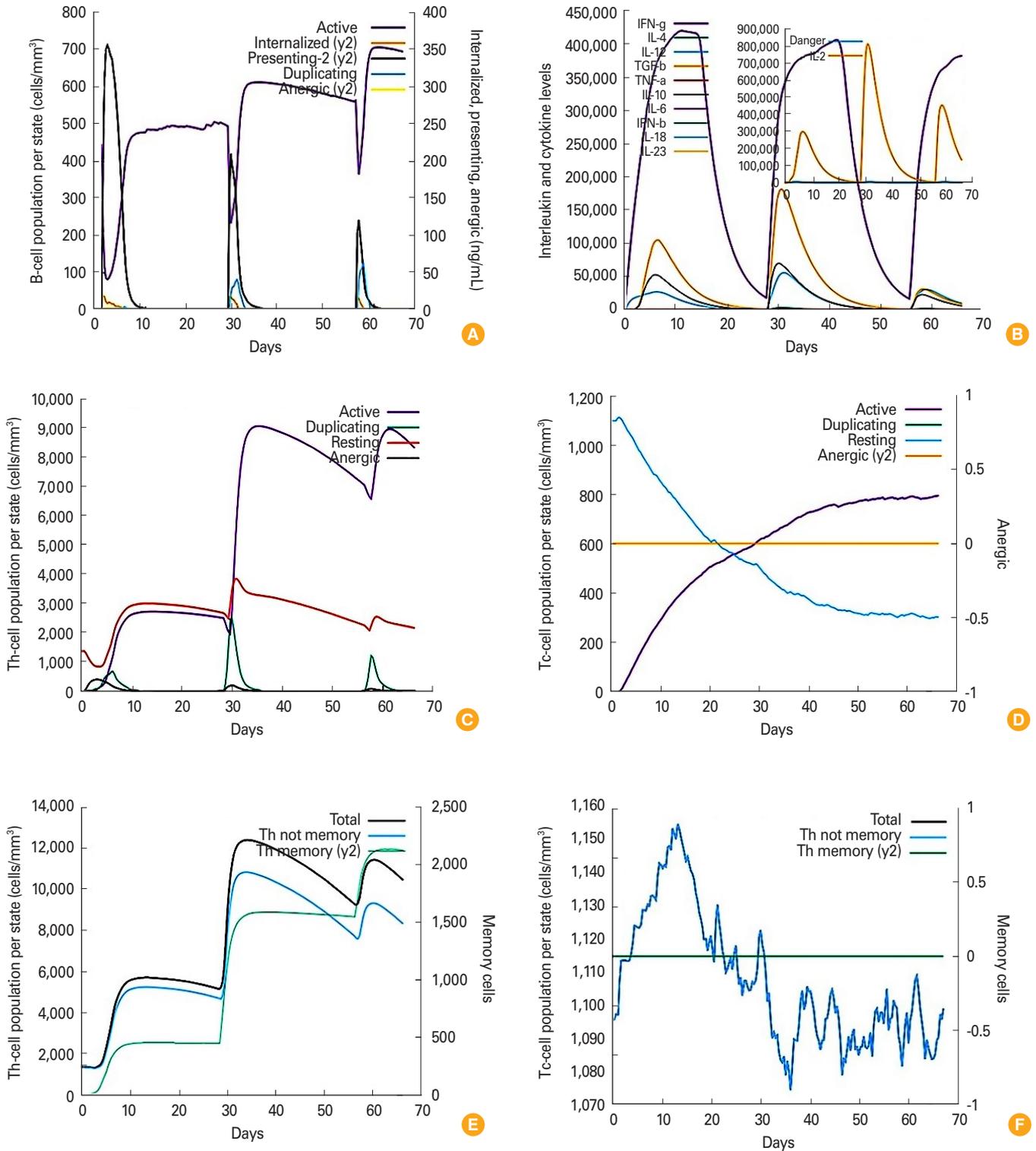


**Fig. 8.** Ramachandran plot for the tertiary structure model of messenger RNA vaccine construct against avian influenza A (mVAIA). Green spots represent residues within the highly favored region. The residues depicted as brown triangles are within the favored region while the residue depicted as a brown circle in the lower right quadrant is within the disallowed region.

**Table 3.** B-cell epitopes showing accessible residues in the tertiary structure model of mVAIA

| Type of B-cell epitope | Residues  | Residues | Score |
|------------------------|---|----------|-------|
| Discontinuous epitope  | A:Q1, A:S2, A:L3, A:L4, A:T5, A:E6, A:V7, A:E8, A:T9, A:P10, A:I11, A:R12, A:N13, A:E14, A:W15, A:G16, A:C17, A:R18, A:G19, A:P20, A:G21, A:P22, A:G23, A:Q24, A:Y25, A:I26, A:K27, A:A28, A:N29, A:S30, A:K31, A:F32, A:I33, A:G34, A:I35, A:T36 | 36       | 0.8   |
| Linear epitope         | QSSLTEVETPIRNEWGCRGPGPGQYIKANSKFIGITG   | 36       | 0.821 |

mVAIA, messenger RNA vaccine construct against avian influenza A.



**Fig. 9.** Immune simulation plots for the secreted protein sequence of messenger RNA vaccine construct against avian influenza A (mVAIA). B-cell population count per entity-state for active, presenting on class-II, internalized the antigen, duplicating, and anergic (A); cytokine and interleukin concentration level with danger signal indicated (B); helper T-cell population counts for active, resting, and duplicating cells (C); cytotoxic T (Tc)-cell counts for active, duplicating, resting, and anergic cells; total and memory T-helper (Th) cell counts (D); and total and memory T-cytotoxic lymphocytes. IFN, interferon; IL, interleukin; TGF, transforming growth factor.

to simulate potential immune responses induced by the vaccine (Fig. 9). After 3 injections of the same antigen with 4 weeks intervals, the populations of activated and duplicating B lymphocytes increased (Fig. 9A). Results of the simulation showed that the antigen can potentially induce cytokines, including interferon-gamma, transforming growth factor-beta, interleukin (IL)-10, and IL-12. The amount of IL-2 induced was also increased but the highest level did not go beyond the inset plot danger signal (Fig. 9B). Both Th (Fig. 9C) and Tc (Fig. 9D) cell populations were activated after the 1st antigen injection. Moreover, activated T lymphocyte populations further increased compared to resting population after the 2nd and the 3rd injection. Populations of memory Th-cell (Fig. 9E) and Tc-cell (Fig. 9F) started to increase even after the 1st administration of the vaccine. The peaks are more dominant relative to the total number of lymphocytes after the 2nd injection and the 3rd injections.

## Discussion

The challenges in antigenic drift and antigenic shift of AIAV warrant avian influenza vaccine reformulation to target multiple antigens across the currently circulating strains that could also confer longer immune protection with just a single dose. Compelled by this demand, an mRNA vaccine construct was conceptualized and strategically designed to incorporate conserved and experimentally validated epitopes of multiple AIAV antigens, including all the necessary elements for its expression (Fig. 2). This study was able to identify non-toxin, conserved linear B-cell and CD8<sup>+</sup> epitopes from the list of experimentally validated AIA epitopes in IEDB. It has been widely accepted that conserved protein sequences have Shannon index  $\leq 2.0$  [28]. Experimentally validated linear B-cell epitope SLLTEVETPIRNEWGCR spans the extravirion region of MP2 protein sequence (1 to 22), making it exposed for paratope binding. One of the most important factors to be considered in excluding predicted B-cell epitopes is the presence of post-translational modification (PTM) sites such as glycosylation sites that may mask the epitopes [29]. In the case of SLLTEVETPIRNEWGCR, cysteine residue (C17) near the C-terminus, is involved in the formation of a disulfide bond with C19 (not part of the epitope). Nevertheless, the presence of this PTM site did not seem to effectively hinder the binding of antibody to the epitope as this linear epitope has been studied for immune reactivity and resulted positive in B-cell assays which included *in vivo* immunization in

chickens and positive qualitative binding in ELISA.

Reported binding energies of MHC class I with peptides can range from -4.1 to -10.7 kcal/mol [30]. Calculated binding energies for the docking of epitopes with chicken MHC class I in this study, range from -28.45 to -40.59 kJ/mol (-6.8 to -9.7 kcal/mol) which suggest favorable complex formation. A study reported that the acceptable binding energy of a peptide and MHC class II complex is lower than -5.8 kcal/mol [31]. The estimated binding energies for chicken MHC class II with TET and PADRE are lower than the  $\Delta G_{\text{bind}}$  of experimental peptide-MHC II complex; thus, favorable binding of chicken BL2 02 with TET/PADRE is plausible. Lastly, it has been established that the dissociation constant (Kd) is inversely proportional to the binding affinity. The Kd values of all the epitope-MHC complexes in this study are  $> 1.0$ . Generally, it has been established that Kd value  $< 1$  indicates that the formation of the complex is more favorable than the reverse reaction. Favorable and stable peptide-MHC complex formation can lead to successful peptide presentation. Overall results suggest that the epitopes investigated in this study can favorably bind to the binding groove of their respective MHC for subsequent induction of immunogenicity. The scores of docking processes suggest high accuracy which imply the reliability of the docked structure models obtained in this study. All these epitopes were included in the proposed mRNA vaccine construct (mVAIA) that can potentially induce humoral immune responses against MP2 and cellular immune responses against HA, MP1, NP, and PAP of AIAV (Fig. 2).

Activation of effector Th cells with the inclusion of the aforementioned universal Th epitopes can subsequently induce Th-dependent B-cell activation and memory B-cell production. However, for mRNA to induce humoral immune response, the linear B-cell and Th cell epitopes must be expressed, secreted, and made available extracellularly for MHC II presentation by antigen presenting cells (APC) in order to activate Th lymphocytes via the exogenous pathway. And this will also facilitate the binding of B-cell epitopes to BCRs for the Th-dependent activation of B-cells. Therefore, the vaccine was designed to contain a PLAT sequence which serves as a signal peptide to improve its secretory expression. Studies have shown that the attachment of PLAT to several proteins has increased the induction of antibody and cellular immune responses in animal models [32]. In addition, it can also take advantage of the cross-presentation mechanisms known for professional APCs for Tc cells presentation through

MHC I [33]. Thus, this strategy can potentially activate both the humoral and the cellular arms of immunity.

The efficacy of an mRNA vaccine is also dependent on its translational efficiency. One approach to increase the protein expression of mVAIA is to prevent its degradation by camouflaging the vaccine as a self mRNA. The use of the host's self-recognized 5' cap 1 avoids immune recognition by intracellular TLRs resulting in increased protein production level [34]. A functional 5' cap structure can be added co-transcriptionally or enzymatically post-transcription. With the aim of further increasing translational efficiency, the sequences of UTRs were derived from the target host to allow efficient recognition and binding of the chicken's inherent translational factors. The Kozak consensus sequence was included in the 5' UTR as it has been demonstrated to play a major role in the initiation of the translation while contributing to the mRNA stability and translation efficiency [16]. Moreover, the poly(A) leader sequence was shown to increase translation efficiency by preventing the formation of secondary structure at the 5' end of the vaccine [16]. Fig. 4 shows that the 5' UTR of mVAIA possesses a single stranded external loop which indicates that it can easily accommodate the translational machinery and facilitate the unwinding of adjacent secondary structures in the initiation site of the sequence. Moreover, the evaluation of the secondary structure of mVAIA demonstrated a highly acceptable thermal stability. Particularly, the secondary structure of mVAIA has a very low  $\Delta G_{\text{fold}}$  value. RNA secondary structures with  $\geq -80$  kcal/mol (334.72 kJ/mol) is regarded as unstable [23]. The MFE can be estimated by the predicted  $\Delta G_{\text{fold}}$  from a random coil RNA to an ordered structure where more negative values indicate greater stability [35]. Assessment of minimum free energy for the RNA structure of mVAIA indicates its thermal stability which can make vaccine storage and administration more favorable especially for countries near the equator. The poly(A) tail sequence was set to 100 nucleotides long, as this is deemed to be an ideal length in synthesizing therapeutic mRNA to further regulate its stability and translational efficiency [36]. The mVAIA sequence can be obtained by *in vitro* transcription of an enzymatically linearized plasmid. Fig. 3 shows the plasmid map encoding all the essential elements for the transcription of mVAIA gene.

Analysis of the translation product showed that there is a very high probability for the adjoined PLAT sequence to be recognized as a signal peptide Sec/SPI as attached to the whole amino acid sequence. This step serves the purpose for

including and positioning PLAT sequence for targeted secretory expression of mVAIA. The protein was also classified as antigenic, slightly basic, slightly hydrophobic as indicated by GRAVY index greater than zero [37], and possesses thermostability over a wide range of temperature as indicated by its instability index. Information on its physicochemical properties can be used as references for future vaccine formulation conditions. *In silico* evaluations of allergenicity and cross-reactivity imply that the secreted protein sequence of mVAIA is less likely to cause allergic reactions and induce cross-reaction with protein sequences of chickens. Generally, these results suggest the safety profile of the candidate vaccine.

It is of paramount importance to consider investigating the accessibility of the adjoined linear B-cell epitope (SLLTEVETPIRNEWGCR) for effective BCR binding. The secondary and tertiary structure analysis of its secreted form showed that SLLTEVETPIRNEWGCR is within the region dominated by random coils (disordered), and it is a solvent-accessible moiety of the vaccine's tertiary structure, as indicated by its high protrusion score. The acceptable quality-assessment scores underline the validity of the tertiary structure model generated and utilized for the vaccine.

Lastly, immune simulation plots generated for the secreted protein sequence of mVAIA showed that the vaccine can potentially induce cytokine secretion, activate immune effectors including B and T lymphocytes, and produce memory Th and Tc lymphocytes even after the 1st injection. The levels of cytokine production, as well as the populations of active and memory lymphocyte, have noticeably increased after the 2nd and the 3rd antigen administration based from the simulation plots. These results further support the potential immunogenicity of mVAIA as a vaccine against AIAV.

In conclusion, the incorporation of conserved epitopes can potentially induce cross-protection against various strains of AIAV. Analysis of its optimized sequence, structure, physicochemical properties, accessibility of adjoined B-cell epitopes, potential toxicity, and cross-reactivity, showed that mVAIA can potentially induce humoral and cellular immune responses, and possess good safety profile as a candidate vaccine. The mVAIA construct is anticipated to be validated *in vitro* and *in vivo* in subsequent studies.

## ORCID

Leana Rich Herrera-Ong <https://orcid.org/0000-0003-2243-2747>

**References**

1. Horwood PF. Avian influenza H5N1: still a pandemic threat? *Microbiol Aust* 2021;42:152-5.
2. Yamamoto Y, Nakamura K, Mase M. Survival of highly pathogenic avian influenza H5N1 virus in tissues derived from experimentally infected chickens. *Appl Environ Microbiol* 2017;83:e00604-17.
3. Beerens N, Heutink R, Harders F, Bossers A, Koch G, Peeters B. Emergence and selection of a highly pathogenic avian influenza H7N3 virus. *J Virol* 2020;94:e01818-9.
4. Taubenberger JK, Kash JC, Morens DM. The 1918 influenza pandemic: 100 years of questions answered and unanswered. *Sci Transl Med* 2019;11:eaau5485.
5. Bellini S, Scaburri A, Colella EM, et al. Epidemiological features of the highly pathogenic avian influenza virus H5N1 in a densely populated area of Lombardy (Italy) during the epidemic season 2021-2022. *Viruses* 2022;14:1890.
6. Wappes J. H5N1 avian flu now affecting more than two thirds of states [Internet]. Minneapolis (MN): Center for Infectious Disease Research and Policy; 2022 [cited 2023 Jan 3]. Available from: <https://www.cidrap.umn.edu/h5n1-avian-flu-now-affecting-more-two-thirds-states>
7. Jocson LM. Bird flu detected in 9 provinces in Luzon, 1 in Mindanao. *BusinessWorld* [Internet]. 2022 Oct 6 [cited 2023 Jan 3]. Available from: <https://www.bworldonline.com/the-nation/2022/10/06/479144/bird-flu-detected-in-9-provinces-in-luzon-1-in-mindanao/>
8. Guyonnet V, Peters AR. Are current avian influenza vaccines a solution for smallholder poultry farmers? *Gates Open Res* 2020;4:122.
9. Jain S, Venkataraman A, Wechsler ME, Peppas NA. Messenger RNA-based vaccines: past, present, and future directions in the context of the COVID-19 pandemic. *Adv Drug Deliv Rev* 2021;179:114000.
10. Yu L, Pan J, Cao G, et al. AIV polyantigen epitope expressed by recombinant baculovirus induces a systemic immune response in chicken and mouse models. *Virol J* 2020;17:121.
11. James CM, Foong YY, Mansfield JP, Fenwick SG, Ellis TM. Use of tetanus toxoid as a differentiating infected from vaccinated animals (DIVA) strategy for sero-surveillance of avian influenza virus vaccination in poultry. *Vaccine* 2007;25:5892-901.
12. Li J, Helal ZH, Karch CP, et al. A self-adjuvanted nanoparticle based vaccine against infectious bronchitis virus. *PLoS One* 2018;13:e0203771.
13. Palmer E, Migalska M, Wise D, Tregaskes C, Kaufman J. Transport studies with the polymorphic TAP molecules in chickens. *Mol Immunol* 2022;150:28.
14. Halabi S, Ghosh M, Stevanovic S, et al. The dominantly expressed class II molecule from a resistant MHC haplotype presents only a few Marek's disease virus peptides by using an unprecedented binding motif. *PLoS Biol* 2021;19:e3001057.
15. Diatrotov ME. Changes in body temperature of small mammals and birds in a few minutes range as reflection of environmental influences. *Bull Exp Biol Med* 2021;171:388-92.
16. Dhungel P, Cao S, Yang Z. The 5'-poly(A) leader of poxvirus mRNA confers a translational advantage that can be achieved in cells with impaired cap-dependent translation. *PLoS Pathog* 2017;13:e1006602.
17. Kaczmarek JC, Kowalski PS, Anderson DG. Advances in the delivery of RNA therapeutics: from concept to clinical reality. *Genome Med* 2017;9:60.
18. Rohner E, Yang R, Foo KS, Goedel A, Chien KR. Unlocking the promise of mRNA therapeutics. *Nat Biotechnol* 2022;40:1586-600.
19. Dimitrov I, Flower DR, Doytchinova I. AllerTOP: a server for in silico prediction of allergens. *BMC Bioinformatics* 2013;14 Suppl 6(Suppl 6):S4.
20. Ponomarenko J, Bui HH, Li W, et al. ElliPro: a new structure-based tool for the prediction of antibody epitopes. *BMC Bioinformatics* 2008;9:514.
21. Castiglione F, Duca K, Jarrah A, Laubenbacher R, Hochberg D, Thorley-Lawson D. Simulating Epstein-Barr virus infection with C-ImmSim. *Bioinformatics* 2007;23:1371-7.
22. Baldazzi V, Castiglione F, Bernaschi M. An enhanced agent based model of the immune system response. *Cell Immunol* 2006;244:77-9.
23. Mohammadin S, Edger PP, Pires JC, Schranz ME. Positionally-conserved but sequence-diverged: identification of long non-coding RNAs in the Brassicaceae and Cleomaceae. *BMC Plant Biol* 2015;15:217.
24. Georgakopoulos-Soares I, Parada GE, Hemberg M. Secondary structures in RNA synthesis, splicing and translation. *Comput Struct Biotechnol J* 2022;20:2871-84.
25. Guruprasad K, Reddy BV, Pandit MW. Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting in vivo stability of a protein from its primary sequence. *Protein Eng* 1990;4:155-61.

26. Hema K, Ahamad S, Joon HK, Pandey R, Gupta D. Atomic resolution homology models and molecular dynamics simulations of Plasmodium falciparum tubulins. *ACS Omega* 2021;6:17510-22.
27. Mora Lagares L, Minovski N, Caballero Alfonso AY, et al. Homology modeling of the human P-glycoprotein (ABCB1) and insights into ligand binding through molecular docking studies. *Int J Mol Sci* 2020;21:4058.
28. Litwin S, Jores R. Shannon information as a measure of amino acid diversity. In: Perelson AS, Weisbuch G, editors. *Theoretical and experimental insights into immunology*. Berlin: Springer-Verlag; 1992. p. 279-87.
29. Lisowska E. Antigenic properties of human glycoporphins: an update. *Adv Exp Med Biol* 2001;491:155-69.
30. Saito Y, Peterson PA, Matsumura M. Quantitation of peptide anchor residue contributions to class I major histocompatibility complex molecule binding. *J Biol Chem* 1993;268:21309-17.
31. Reich Z, Altman JD, Boniface JJ, et al. Stability of empty and peptide-loaded class II major histocompatibility complex molecules at neutral and endosomal pH: comparison to class I proteins. *Proc Natl Acad Sci U S A* 1997; 94:2495-500.
32. Hanke T, Ondondo B, Abdul-Jawad S, Roshorm Y, Bridgeman A. Vector delivery-dependant effect of human tissue plasminogen activator signal peptide on vaccine induction of T cells. *J HIV AIDS* 2016;2:1-8.
33. Reche PA. Potential cross-reactive immunity to SARS-CoV-2 from common human pathogens and vaccines. *Front Immunol* 2020;11:586984.
34. Drazkowska K, Tomecki R, Warminski M, et al. 2'-O-Methylation of the second transcribed nucleotide within the mRNA 5' cap impacts the protein production level in a cell-specific manner and contributes to RNA immune evasion. *Nucleic Acids Res* 2022;50:9051-71.
35. Andrews RJ, Baber L, Moss WN. Mapping the RNA structural landscape of viral genomes. *Methods* 2020;183:57-67.
36. Kim SC, Sekhon SS, Shin WR, et al. Modifications of mRNA vaccine structural elements for improving mRNA stability and translation efficiency. *Mol Cell Toxicol* 2022;18:1-8.
37. Magdeldin S, Yoshida Y, Li H, et al. Murine colon proteome and characterization of the protein pathways. *BioData Min* 2012;5:11.