



OPEN

## Immunoinformatics, molecular docking and dynamics simulation approaches unveil a multi epitope-based potent peptide vaccine candidate against avian leukosis virus

Siham O. Elshafei<sup>1</sup>, Nuha A. Mahmoud<sup>1</sup> & Yassir A. Almofti<sup>2</sup>✉

Lymphoid leukosis is a poultry neoplastic disease caused by avian leukosis virus (ALV) and is characterized by high morbidity and variable mortality rates in chicks. Currently, no effective treatment and vaccination is the only means to control it. This study exploited the immunoinformatics approaches to construct multi-epitope vaccine against ALV. ABCpred and IEDB servers were used to predict B and T lymphocytes epitopes from the viral proteins, respectively. Antigenicity, allergenicity and toxicity of the epitopes were assessed and used to construct the vaccine with suitable adjuvant and linkers. Secondary and tertiary structures of the vaccine were predicted, refined and validated. Structural errors, solubility, stability, immune simulation, dynamic simulation, docking and in silico cloning were also evaluated. The constructed vaccine was hydrophilic, antigenic and non-allergenic. Ramchandran plot showed most of the residues in the favored and additional allowed regions. ProsA server showed no errors in the vaccine structure. Immune simulation showed significant immunoglobulins and cytokines levels. Stability was enhanced by disulfide engineering and molecular dynamic simulation. Docking of the vaccine with chicken's TLR7 revealed competent binding energies. The vaccine was cloned in pET-30a(+) vector and efficiently expressed in *Escherichia coli*. This study provided a potent peptide vaccine that could assist in tailoring a rapid and cost-effective vaccine that helps to combat ALV. However, experimental validation is required to assess the vaccine efficiency.

Avian leukosis virus (ALV) is one of the most prevalent retroviruses causing neoplasms in chickens<sup>1,2</sup>. ALV related to the genus *Alpharetrovirus* and subfamily *Orthoretrovirinae*<sup>3</sup>. ALVs were shown causing avian leukemia/sarcoma in birds with major economic losses in the global poultry industry<sup>1,4</sup>. ALV causes high morbidity rates than mortality rates in the infected flocks which negatively impacts hatchability and egg productivity<sup>5-7</sup>. In addition to the neoplasia, other factors contributed to mortality, including delayed maturation and diminished fertility rate, which have serious economic implications<sup>4</sup>. The co-infections of the virulent ALV isolates with reticuloendotheliosis virus (REV) and marek's disease virus (MDV) are the main factor for prevalence of avian neoplastic diseases in China that hampering the future control of disease<sup>3</sup>. Therefore the affected flocks showed lower egg production and hemorrhages in the skin surrounding the feather follicles and phalanges<sup>8,9</sup>. As a result of loss of appetite, chickens characterized by weight loss, abnormal feathers, pale combs and wattles, depression, paralysis and finally death<sup>8</sup>. Postmortem studies revealed enlarged livers, spleens, hearts, kidneys, and congested lungs in the affected chickens<sup>7,9,10</sup>.

Horizontal, vertical, and genetic transmissions are the main three ways of ALV transmission<sup>8</sup>. The horizontal transmission occurs through direct contact with infected chickens and materials. In vertical infection, the disease is transmitted from hens to their off-springs via the eggs. Genetic transmission is through endogenous

<sup>1</sup>Department of Biochemistry, Faculty of Medicine and Surgery, National University, Khartoum, Sudan. <sup>2</sup>Department of Molecular Biology and Bioinformatics, College of Veterinary Medicine, University of Bahri, P.O. Box 1660, Khartoum, Sudan. ✉email: yamofti99@gmail.com

retroviruses occurred through a mendelian manner<sup>8</sup>. Another possible transmission route is through contaminated live and attenuated vaccines as some ALV strains were isolated from contaminated live vaccines<sup>11</sup>.

ALV is a member of the leukosis/sarcoma (L/S) group of avian retroviruses. It is an RNA virus family that uses the enzyme reverse transcriptase which originates the generation of the DNA provirus that is combined with the host genome during viral replication<sup>12</sup>. In terms of structure, ALV is a simple virus and has three chief genes located from its 5'-to-3' end of its genome. These genes are gag/pro, pol and env genes, which code for the antigens and proteases that distinguish a different group of viron, as well as, the reverse transcriptase enzyme and the glycoprotein or the envelope protein<sup>12</sup>. The genomic sequence associated with an ordinance of the viral replications that flank these structural genes in the DNA provirus, forming long terminal repeats (LTRs) for the virus with promoters and enhancers<sup>13</sup>. Based on their viral envelope characteristics, chicken leukosis virus envelope subgroups are categorized into seven subgroups identified as A, B, C, D, E, J and K. The exogenous subgroups A, B, C, D, and J are highly pathogenic, while those of subgroup E viruses are endogenous and less pathogenic<sup>10,14</sup>. Virus subgroups are mainly differentiated by the receptors they use, which determine their host range and tropism. The highly pathogenic ALV-K subgroup has recently been detected in China based solely on its genomic sequence, posing a major threat to Asia's poultry industry<sup>14,15</sup>. In ALV genomes, subgroup differences can be primarily detected in the env gene, which binds to cell-specific receptors and dictates viral pathogenicity and tropism<sup>16</sup>.

Despite more than a century of ALV-related research, no vaccine and no specific treatment exist to prevent chickens from the disease. Limited anti-ALV vaccines are being developed. Among them four targeting ALV strain J, one targeting strain B, and another one targeting strain A<sup>17</sup>. Also, a limited protection and less immunogenicity were observed in some anti-ALV vaccine trials<sup>17</sup>. However defeating and preventing the ALV disease from spreading has been accomplished by eliminating affected birds, managing hatcheries and farms and implementing biosecurity measures<sup>1,4</sup>.

With the proliferation of big genomic and proteomic data and computing power, a successful vaccine could be developed by accessing vast information about the virus structure, protein structures, and predicted epitopes. A combination of computational methods and immunoinformatics had formed the bases for the development of multi-epitopes vaccines<sup>18,19</sup>. In reverse vaccinology, which analyzes a pathogen's genome and identifies its prospective antigenic proteins, multi-epitope vaccines are produced in a shorter period of time and at a lower cost, with safety and highest therapeutic efficiency and minimal adverse reactions<sup>18,19</sup>. In addition, to the boosting of immunogenicity, vaccines with multi-epitopes may improve immune responses because they contain epitopes representative of different target genes<sup>20</sup>. Moreover, this novel computational science can restrict the use of animal models and replaces some clinical trials<sup>21</sup>. Therefore, designing and developing multi-epitope-based peptide vaccines via computational science holds great promise for controlling and preventing the occurrence of infectious diseases in the future. Despite the advances in the field of in silico vaccines formulation, some drawbacks were emerged. For instance, the in silico approaches cannot predict lipids or polysaccharides as they are significant active constituents in proteins<sup>22</sup>. Also predictions of optimum antigen-processing sites are not well addressed and clarified by many epitope prediction tools<sup>23</sup>. On the other hand, common traditional vaccines might use inactivated or attenuated forms of microorganism, which might provide adverse effects and autoimmune diseases in infected hosts. Also the development of traditional vaccines candidate such as chimeric virus vaccines, live attenuated vaccines, virus-like particles (VLPs), subunit vaccines and nucleic acid vaccines is complex process, expensive, time consuming, need complicated in vitro and in vivo studies for evaluating their potential efficacy<sup>24,25</sup>.

Hence, this research targeted the use of immunoinformatics tools to predict multi epitopes vaccine that could efficiently combat the ALV infection. The proteome of ALV was investigated, evaluated and analyzed for epitopes prediction. The results demonstrated that the constructed vaccine has successfully comprised all the most favorable factors to be a potential vaccine candidate, which further requires experimental validation. This study provided good theoretical basis essential for developing a new vaccine strategy and extending the current knowledge in controlling ALV.

## Materials and methods

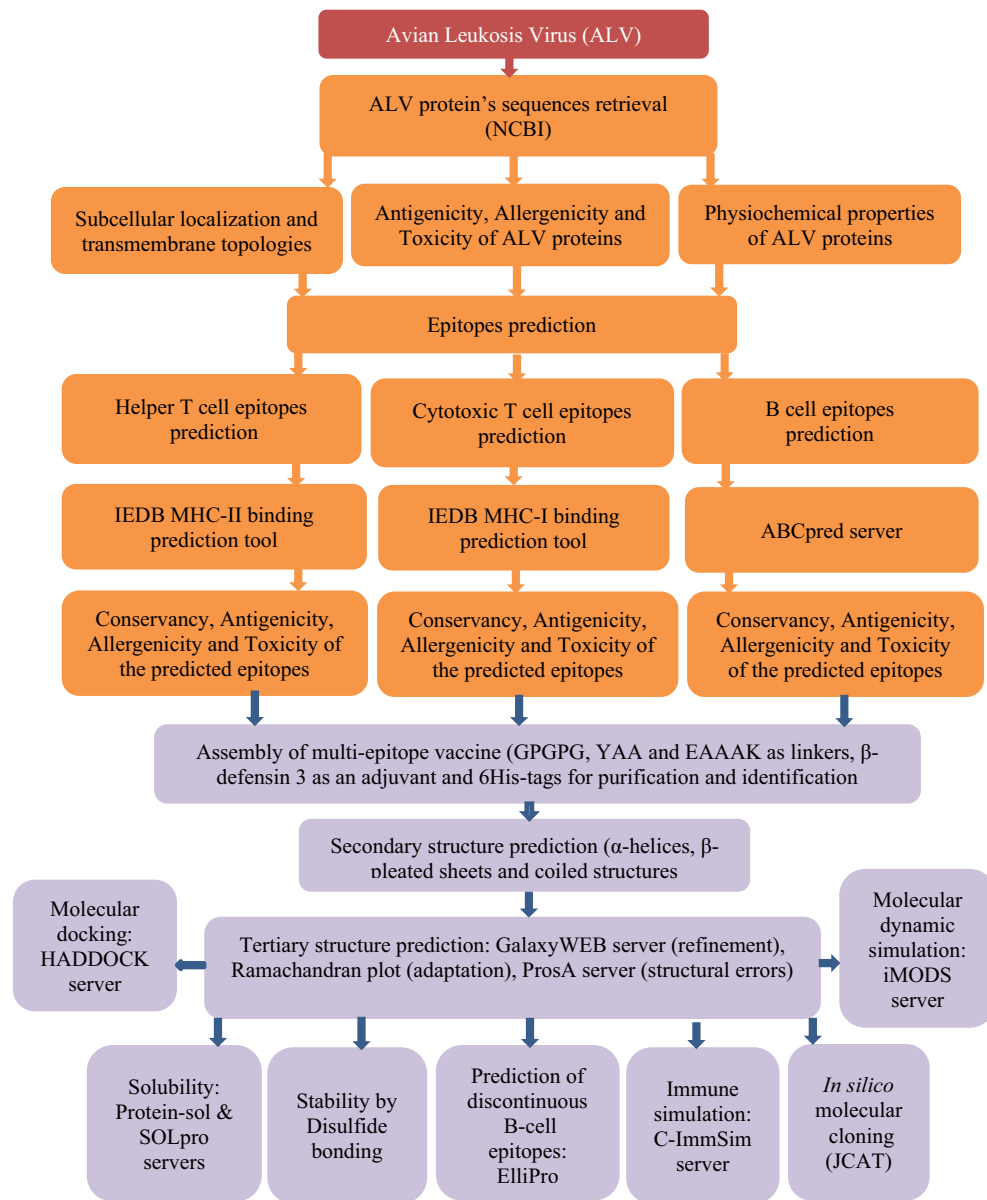
The immunoinformatics steps for the in silico vaccine design were visualized in the flow chart presented in Fig. 1.

### ALV protein's sequences retrieval

The ALV demonstrated three proteins: polymerase protein, envelope protein, and transacting factor protein with the following accession numbers NP\_040550.1, NP\_040548.1, and NP\_040549.1, respectively. The sequences of these three proteins were retrieved from the National Center for Biotechnology Information (NCBI) at (<https://www.ncbi.nlm.nih.gov/protein>)<sup>26</sup>.

### Physicochemical properties, antigenicity, allergenicity, and toxicity of the retrieved proteins

A physicochemical profile of each retrieved protein was assessed using the Protparam server (<http://web.expasy.org/protparam/>)<sup>27</sup>. The potent antigenicity was determined using the Vaxijen v2.0 server (<http://www.ddg-pharmfac.net/vaxijen/>) based on a default viral threshold of 0.4 for each antigenic protein<sup>28,29</sup>. Allergenicity and toxicity of the proteins were assessed using the AllerTOP server (<http://www.ddg-pharmfac.net/AllerTop/>)<sup>30</sup> and the ToxinPred server (<http://crdd.osdd.net/raghava/toxinpred/>)<sup>31</sup>, respectively. The same servers were later used to assess the physicochemical properties, antigenicity, allergenicity and toxicity properties of the constructed vaccine.



**Figure 1.** Schematic flowchart providing the overall steps used for designing the ALV multi-epitope based peptide vaccine.

### Subcellular localization and transmembrane topologies of virus proteins

Subcellular localization of viral proteins is considered as an important clue to the function of the immune cells and judging the potential efficacy of vaccine targets<sup>32</sup>. In addition, surface-localized proteins are among the best candidates for the recombinant vaccine, since they are the first molecular patterns of pathogens that contacted by the host immune system<sup>32</sup>. For the detection of the viral protein subcellular localization, the Phobius server (<https://phobius.sbc.su.se/index.html>) was used<sup>33</sup>. The server provided a combination of a transmembrane topology (TMHs) and a signal peptide predictor.

### Epitopes prediction and conservancy

A total of 50, 13, and 3 strains sequences were retrieved for the polymerase, envelope, and transacting factor protein, respectively. These strains were used for epitopes conservancy and were presented in Table 1. BioEdit program version 7.2.5, is a multiple sequence alignment (MSA) tool, was used to align each protein strains sequences<sup>34</sup>. The analysis of the aligned sequences was conducted in order to identify the conserved epitopes that effectively act against B and T lymphocytes. Epitopes that had 100% conservancy (no mutations) among the strains were selected for further analysis, while non-conserved epitopes were excluded.

Polymerase protein	Polymerase protein	Polymerase protein	Polymerase protein	Envelope protein	Transacting protein
*NP_040550.1	AKP18447.1	ADO34849.1	BAC78443.1	*NP_040548.1	*NP_040549.1
AKP18453.1	BAM15580.1	ADO34846.1	QEJ82123.1	AAA91268.1	AAA91267.1
AKP18448.1	BAM68487.1	ADO34843.1	QEG03935.1	AIY32623.1	Q04221.1
AIY32622.1	BAM68485.1	AHI49963.1	AWO14323.1	BAM68473.1	
ADP21278.1	BAM68481.1	AFG30999.1	ARH12433.1	AYN55358.1	
ADP21275.1	BAM68477.1	ASM94215.1	ARH12430.1	ABX39002.1	
BAK64409.1	BAM68472.1	AMP18914.1	AAA91269.1	ABX38999.1	
AYN55373.1	BAM68470.1	AEF97638.1	ABX39004.1	ABX38996.1	
AYN55371.1	BAL70360.1	AHC94804.1	ABX39001.1	ADN85628.1	
AYN55368.1	BAL70357.1	AIM55740.1	ABX38998.1	ADO34853.1	
AYN55365.1	BAL70354.1	BAA22090.1	ABC94886.1	ABC94884.1	
AYN55359.1	BAL70351.1	BAL70337.1		AYC62336.1	
ANW72069.1	AIL52058.1	BAK64246.1		WGM49020.1	

**Table 1.** The total number of the retrieved strains of the polymerase, envelope, and transacting protein of ALV with their accession numbers. \*Reference sequence.

### B cell epitopes prediction

To identify the B cell epitopes from the three ALV proteins, the ABCpred server ([https://webs.iitd.edu.in/ragha/abcpred/ABC\\_submission.html](https://webs.iitd.edu.in/ragha/abcpred/ABC_submission.html)) was utilized<sup>35,36</sup>. The reference sequence of each protein was used as an input into ABCpred server. The default thresholds (0.51), as well as 12mers epitope length were used for the prediction strategy. VaxiJen v2.0, AllerTOP, and ToxinPred servers were further used to assess the antigenic, allergenic and toxic potentiality of the conserved B cell epitopes, respectively.

### T cell epitopes prediction

Based on Immune Epitope Database (IEDB) analysis resources at (<https://www.iedb.org/>), different T cell epitope prediction tools were analyzed<sup>37,38</sup>. The reference sequence was used as an input for each protein analysis. The data for epitopes that interacted with the major histo-compatibility complex class I and II (MHC-I and MHC-II) is not yet organized in the IEDB resources for chicken alleles. Accordingly, the human alleles were exploited to predict epitopes from the ALV-retrieved proteins interacting with T cell epitope as previously described<sup>39,40</sup>.

#### Cytotoxic T cell epitopes prediction

The IEDB prediction method at (<http://tools.iedb.org/mhci/>) provided a number of MHC-I binding prediction methods. In this study, the prediction of the MHC-I interacted alleles were obtained by Artificial Neural Network, NetMHC (ANN)<sup>41</sup>. The human reference alleles sets (HLA-A, HLA-B, and HLA-C) were used for the prediction process. Conserved epitopes with a score equal to or less than 1 ( $\leq 1$ ) percentile rank with nine amino acids length bound to alleles were only analyzed. The conserved cytotoxic T-cell epitopes were further assessed for antigenicity, allergenicity and toxicity predictions.

#### Helper T cell epitopes prediction

The IEDB MHC-II binding prediction tool (<http://tools.iedb.org/mhcii/>) was used to investigate the reference sequence of ALV proteins for epitope prediction against MHC-II<sup>37,38</sup>. The human alleles reference sets (HLA-DP, HLA-DQ, and HLA-DR) were employed to search for promising epitopes. The analysis comprises the Neural Networks Align method, NetMHCII, version 2.2 (NN-align)<sup>37,38,41</sup>. The approach was used to find potential epitopes having a percentile rank score equal to or less than 10 ( $\leq 10$ ). The core sequence and peptide lengths were set to 9 and 18 amino acids, respectively. The antigenic, allergenic, and toxic evaluation of the conserved helper T cell epitopes was carried out using VaxiJen v2.0, AllerTOP, and ToxinPred servers, respectively.

### Assembly of the multi-epitope vaccine

The primary assembly of the vaccine sequence was accomplished fusing the B and T cells predicted epitopes that demonstrate conservancy, antigenicity score more than 1 and were shown to be non-allergenic and non-toxic. The elected B and T helper epitopes were fused by the GPGPG linkers while the T cytotoxic epitopes were fused by the YAA linkers<sup>42</sup>. The 5'-amino terminal of the vaccine was supported by the  $\beta$ -defensin 3 (Q5U7J2) as an adjuvant after separation with the EAAAK linker. Moreover, the sequence was provided with the 6His-tag for purification and identification of the vaccine upon expression<sup>42-45</sup>.

### Secondary and tertiary structures prediction of the vaccine

Predicting the secondary structure of the vaccine is available for free on the raptor X server (<http://raptorx.uchicago.edu/>)<sup>46</sup>. The secondary structure (SS), disorder regions (DISO), and solvent accessibility (ACC) played crucial roles in predicting  $\alpha$ -helices,  $\beta$ -pleated sheets, and coiled structures, respectively. For the tertiary structure prediction, the vaccine sequence was submitted to the same raptor X server<sup>46</sup>. The obtained results were received in the form of a PDB file that was further used for vaccine sequence refinement and adaptation.

### Vaccine tertiary structure refinement and adaptation

Refinement of the vaccine's tertiary structure was carried out using the GalaxyWEB server (<http://galaxy.seoklab.org/>)<sup>47,48</sup>. The structured model is made more dynamic and effective by utilizing the side chains, establishing them and repacking these side chains to achieve relaxation of the structured model<sup>47,49</sup>. As a result, the refinement and the structure's physical quality were improved. In the SAVES server, the Ramachandran plot (<https://saves.mbi.ucla.edu/results?job=869538&p=procheck>) was used to analyze the model validation<sup>50</sup>, and the ProSA server (<https://prosa.services.came.sbg.ac.at/prosa.php>) was used to analyze the errors in the protein structure<sup>51,52</sup>.

### Determination of the solubility of the vaccine

Protein-sol server (<https://protein-sol.manchester.ac.uk/>) consists of a set of predictive algorithms and a theoretical computation was used to measure the solubility of the vaccine protein<sup>53</sup>. The vaccine solubility was assessed in comparison with *Escherichia coli* proteins, which have an estimated solubility of 0.45. Thus, proteins with solubility score greater than 0.45 are therefore considered soluble<sup>53</sup>. To confirm the solubility of the vaccine, the SOLpro server (<http://scratch.proteomics.ics.uci.edu/>) was further employed<sup>54</sup>. For a protein to be soluble in the SOLpro server the probability should be  $\geq 0.5$  and insoluble protein scores  $< 0.5$ .

### Determination of the stability of the vaccine

The disulfide bonding in a given protein between the cysteine residues plays an important role in strengthening of the protein's geometric conformation and enhances its extensive stability<sup>55</sup>. Disulfide-by-Design 2.0 (DbD2) (<http://cptweb.cpt.wayne.edu/DbD2/>) is a web-based tool that facilitates designing disulfide bonds in vaccine construct by substituting particular amino acid with cysteine in high-mobility and unstable regions of proteins<sup>55</sup>. This was followed by formation of disulfide bonds between cysteine residues. The parameters such as the intra-chain, inter-chain and build  $C_{\beta}$  for Gly were chosen. The angle ( $-87^{\circ}$  or  $+97^{\circ} \pm$ ) was set to 30 and  $C_{\alpha}-C_{\beta}-S_{\gamma}$  angles ( $114.6^{\circ} \pm$ ) was set to 10 for proper prediction of the bonds<sup>55</sup>.

### Immune simulation

For mimicry of the immune response and immunogenicity of the ALV vaccine in the host, C-ImmSim server (<https://kraken.iac.rm.cnr.it/C-IMMSIM/>) was used<sup>56</sup>. Two injections were given with time step set at 1 and 90 (the server provided each time step as 8 h while the time step 1 represents the injection at time = zero). The other simulation parameters were set to default. The measure of diversity (Simpson index, D) was interpreted from the plot<sup>56</sup>.

### Molecular dynamic simulation (MD)

iMODS server (<https://chaconlab.org/multiscale-simulations/imod>) was used to analyze the collective motions of protein vaccine<sup>57,58</sup>. A normal modes analysis (NMA) in internal coordinates is conducted by the server to determine the stability of the vaccine protein. This server structured the dynamics of the protein complex and provided various results data, such as deformability, eigenvalues, B-factors, variance maps, co-variances, elastic networks in the atoms, and residue indexes in terms of magnitude and direction<sup>57,58</sup>.

### Prediction of discontinuous B-cell epitopes

The ElliPro in the IEDB (<http://tools.iedb.org/ellipro/>) was used to predict the discontinuous B cell epitopes<sup>59</sup>. ElliPro tool predicts discontinuous and linear antibody epitopes based on the protein 3D structure. The prediction method was based on the default parameters of the sever<sup>59</sup>. For instance, the minimum score and the maximum distance (Angstrom) of the selected epitopes prediction were set to 0.5 and 6, respectively.

### Active sites detection in the vaccine structure

Searching for a ligand-binding region on a protein is an essential step prior to molecular docking process. The process primarily based on multiple factors such as detection of hydrophobic or hydrophilic interactions, salt bridges and electrostatic and hydrogen bonding interactions. The computed atlas of surface topography of proteins (CASTp 3.0) website (<http://sts.bioe.uic.edu/castp/index.html?3igg>) was exploited to determine the vaccine's active regions<sup>60,61</sup>. The default probe radius of 1.4 Å was used.

### Molecular docking of the vaccine protein with chicken TLR7

Protein-protein interaction is essential for functioning of many biological molecules<sup>62</sup>. Analyzing the complex structures formed between these molecules is of great importance to assess the molecular interactions or the affinities between these molecules. Toll-like receptors (TLRs) are considered as recognition receptors that play a paramount role in recognition of pathogen. In birds, there are ten genes encoding for TLRs, among them, TLR7 was chosen for the docking with the vaccine construct since it is a viral-sensing TLR<sup>63</sup>. Thus the designed vaccine was docked against the chicken TLR7 using the HADDOCK 2.4 server (<https://www.bonvinlab.org/software/haddock2.4/>)<sup>62</sup>. Refinement interface in HADDOCK server was used to provide the accurate cluster. PRODIGY web server (<https://wenmr.science.uu.nl/prodigy/>)<sup>64,65</sup> was used to calculate the binding affinities of the best chosen clusters at 25 °C. Finally, the interaction between the vaccine and the chicken TLR7 was visualized by PDBsum server (<https://www.ebi.ac.uk/thorntonsrv/databases/pdbsum/Generate.html>)<sup>66</sup>.

### In silico molecular cloning and codon adaptation

The in silico cloning ensures that a particular host would express the vaccine protein upon cloning in suitable vector<sup>67</sup>. To facilitate successful cloning, optimization process and cloning of the vaccine construct in the



expression vector were performed. The optimization comprises the elimination of different restriction enzymes cleavage sites, prokaryotic ribosomal binding sites, and rho-independent terminators of transcription in the sequence of the vaccine<sup>67</sup>. A reverse translation of the vaccine protein sequence into a DNA sequence was performed with the Java Codon Adaptation Tool (JCAT) (<http://www.prodoric.de/JCat>) because cloning uses DNA rather than proteins<sup>67</sup>. The codon adaptation index and the GC content were in ranges of 0.8–1.0 and 30–70%, respectively. The sequences of the restriction enzymes Xho1 (5-CTCGAG-3) and BamHI (5-GGATCC-3) were added at the 5' and the 3' ends of the DNA, respectively. A restriction cloning module from SnapGene (<https://www.snapgene.com/>)<sup>67</sup> was used to clone the DNA sequence located between the restriction sites of BamHI and Xho1 in the pET-30a(+) vector.

## Results

### Characteristics of the virus proteome

Polymerase, envelope, and transacting factor proteins from the ALV were retrieved from the NCBI database. These three proteins were found to be stable and hydrophilic using the ProtParam server. The VaxiJen server was used to determine and prove their antigenicity. The three proteins were used as inputs to predict B and T cell epitopes for designing the vaccine against ALV. All the physical and chemical features of the three proteins were provided in Table 2.

### Multiple sequence alignment and epitopes conservancy

The ClustalW program provided in the Bioedit tool was used for multiple sequence alignment (MSA) of all retrieved strains. MSA was exploited to search for conserved epitopes among the retrieved strains from polymerase, envelope, and transacting factor proteins. Epitopes length that was not broken by mutated amino acids from other strains is considered conserved epitope. During the MSA, the retrieved strain sequences demonstrated high epitopes conservancy.

### Linear B-cell epitopes prediction

The ABCpred server received the reference sequences from each protein. In the server, a trained recurrent neural network provided the predicted B-cell epitopes based on their scores. Generally, an epitope passing the threshold of 0.51 is more likely to have a higher peptide score. Based on the ABCpred server, 39, 29, and 10 epitopes were predicted from the polymerase, envelope, and transacting factor proteins, respectively. After assessing the antigenicity, allergenicity and the toxicity of the predicted epitopes from each protein, 11, 10 and 6 epitopes from polymerase, envelope and transacting proteins were chosen as B cell epitopes, respectively. These epitopes were provided in Table 3.

### Cytotoxic T lymphocyte epitopes prediction

Based on the reference sequences of polymerase, envelope, and transacting factor, multiple epitopes were predicted against human alleles (HLA-A, HLA -B, HLA-C) using IEDB MHC-1 binding prediction tools. Antigenic, allergenic, and toxic effects were then assessed for the predicted epitopes. A total of 6, 11, and 15 epitopes were obtained from the polymerase, envelope, and transacting factor proteins, respectively, and were elected as T cytotoxic cell epitopes due to their high antigenicity scores, non-allergenicity, non-toxicity and the allelic interactions. These epitopes were provided in Table 4.

### Helper T lymphocyte epitopes prediction

The reference sequence of each of the three proteins (polymerase, envelope, and transacting factor) was analyzed against the human alleles (HLA-DR, DQ, DP) using IEDB MHC-II binding prediction tools with a percentile rank of ( $\leq 10$ ). A vast amount of epitopes were predicted from the three proteins. The predicted epitopes were analyzed for antigenic, allergenic, and toxic outcomes. A total of 21, 6, and 7 epitopes were obtained from the polymerase, envelope, and transacting factor proteins, respectively. They were elected as T helper cell epitopes due to their high antigenicity scores, non-allergenicity, non-toxicity, and allelic interactions. These epitopes were provided in Table 5.

Viral protein	Accession number	Molecular weight (Kilodalton)	Instability index#	Aliphatic index	Theoretical pI	No amino acids	Extinction Coefficient	GRAVY†	Vaxijen antigenicity\$	Subcellular localization and TMHs
Polymerase partial	NP_040550.1	98.630	38.85	92.01	8.17	895	167,410	-0.157	0.5453	NC with no TMH
Envelope protein	NP_040548.1	66.444	41.90	88.96	8.09	606	92,370	-0.104	0.4782	C with 1TMH
transacting factor @	NP_040549.1	14.098	72.84	76.91	12.05	123	38,500	-0.754	0.4504	NC with no TMH

**Table 2.** Physical and chemical properties, antigenicity, and number of the predicted transmembrane helices of ALV proteins. *THMs* Transmembrane helices, *NC* not cytoplasmic, *C* cytoplasmic. @The protein contains no Phe, Tyr, and Asp. #instability index < 40 considered the protein stable; †GRAVY negative sign indicated the protein is hydrophilic; \$the threshold for the Vaxijen antigenicity is 0.4 and the three proteins were shown to be non-allergic. *THMs*: Transmembrane helices. *NC*: not cytoplasmic. *C*: cytoplasmic.

Protein	Sequence	Start	Score*	Vaxijen antigenicity#
Polymerase protein	VEKELQLGHIEP	38	0.63	1.9908
	VNAKLVPGAVQ	77	0.51	1.746
	WNLDMKMAWREI	291	0.73	1.5822
	AGVNPRGLGPLQ	620	0.62	1.5003
	IARPLHVSJKVR	425	0.74	1.381
	IWQTDFTLEPRM	632	0.66	1.2853
	TLEPRMAPRSWL	638	0.59	1.2682
	PLMVLDLKDCFF	102	0.78	1.2436
	LDLKDCFFSIPL	106	0.7	1.2332
	VLEPLRLKHPSL	164	0.84	1.1143
	KACNISMQQARE	593	0.8	1.1041
	Envelope protein	VGFRPQGVWPWYL	193	0.72
ACWSVKQANLTT		445	0.83	1.6310
LTGYPGKTSKGD		9	0.84	1.4026
KTPLLPTRVNYI		35	0.65	1.3434
ISGITGGCVGFR		185	0.64	1.3339
YNCSQVGRQYRC		256	0.66	1.3537
LWITWANRTGQT		73	0.6	1.3105
TGIRKRKSTSHL		395	0.68	1.1582
LPTRVNYILIG		39	0.6	1.0656
PGKTSKKSKEK		13	0.61	1.0458
Transacting protein		LLMWGVVILPW	78	0.82
	RNQQRSGESTKG	19	0.8	0.9436
	LPQVRSASRNQQ	11	0.78	0.8661
	GVVILPWGWE	82	0.73	0.8582
	PQPRLLLLMWG	71	0.67	0.7729
	KPLAHPAISAEQ	53	0.62	0.4698

**Table 3.** The predicted B cell epitopes and their antigenicity scores. \*The default score of the ABCpred server was 0.51 and the length of the predicted epitopes was 12mers. #The Vaxijen server antigenicity threshold was 0.4. All the tested epitopes were nonallergic and nontoxic.

### Structure of the assembled vaccine

The entire number of predicted B cell, T cytotoxic, and T helper epitopes from the three proteins of ALV were used in the construction of the vaccine. Adjuvant, linkers, and 6-His-tags were also embedded in the final structure of the vaccine. Thus the final vaccine structure comprised 738 amino acids. The antigenicity score of the assembled vaccine was 0.8535 when examined in the Vaxijen server. Also, the vaccine protein was non-allergic in the AllerTOP server.

### Physicochemical properties of the assembled vaccine

ProtParam server was used to examine the physicochemical properties of the assembled vaccine. The predicted vaccine weighed 77.121 kilo Dalton (kd) and possessed a theoretical isoelectric point of 9.81, indicating the proposed vaccine had an alkaline pH. Negatively and positively charged residues were 33 and 79 respectively. The Extinction coefficient at 280 nm measured in water was shown to be 132,125 assuming all pairs of Cys residues forming cystines. The instability index score (II) was 38.24, indicating a stable vaccine protein, while the aliphatic index score was 78.73, indicating a hydrophilic vaccine. The grand average water affinity was -0.130, suggesting a hydrophilic vaccine.

### Secondary and tertiary structures prediction of the assembled vaccine

The SS3, ACC, and DISO for the secondary structure were predicted using the Raptor X server. The SS3 showed 23%, 15%, and 61% of the residues as  $\alpha$ -helix,  $\beta$ -sheets and coiled, respectively. The ACC provided 49% as exposed residues, 21% as medium residues and 29% as buried residues. The DISO (disordered predicted regions) was 43 (5%). Figure 2 showed the primary sequence, the tertiary and the refined structures of the vaccine construct.

### Vaccine tertiary structure refinement and validation

The vaccine's stability was assessed via the Ramachandran plot after refinement. In the plot, 90.9% of residues were located in the most favored region. While regions of additional allowed, generously allowed, and disallowed comprised residues of 6.1%, 1.9%, and 1.0%, respectively (Fig. 3a). The ProSA server provided a Z score of -5.68 demonstrating a favorable model structure (Fig. 3b).

Protein	Epitopes	Alleles	Start	End	PR*	Vaxijen Antigenicity <sup>#</sup>
polymerase protein	GHIEPSLSC	HLA-B*15:09	45	53	0.16	1.9236
		HLA-B*42:01		0.37		
	HIGPRALSK	HLA-A*03:01	585	593	0.23	1.6454
		HLA-B*08:02		0.30		
		HLA-B*27:20		0.86		
	LEARAVAMA	HLA-B*45:01	483	491	0.17	1.5777
	VLDLKDCCF	HLA-A*01:01	105	113	0.90	1.4679
		HLA-A*02:11		0.38		
	AMAAVLHVR	HLA-A*02:50	539	547	0.65	1.3796
	RCEQGAIGV	HLA-C*05:01	326	334	0.27	1.2224
HLA-A*02:11			0.66			
Envelope protein	RLACWSVKQ	HLA-B*08:03	443	451	0.95	2.3391
	CWSVKQANL	HLA-A*24:03	446	454	0.59	2.1629
	WSVKQANLT	HLA-C*15:02	447	455	0.73	1.7785
	LQNRAAIDF	HLA-A*66:01	474	482	0.73	1.3636
	DFCLSTQSA	HLA-B*14:02	85	93	0.37	1.1516
	AAIDFLLA	HLA-A*02:06	478	486	0.15	1.1476
	SVKQANLTT	HLA-A*30:01	448	456	0.71	1.0535
	QNRAAIDFL	HLA-B*08:03	475	483	0.34	0.9233
	AVLQNRAAI	HLA-A*32:07	472	480	0.24	0.8515
	YLGKLTMLA	HLA-A*02:03	371	379	0.15	0.8309
	VVTADRHNL	HLA-C*07:01	230	238	0.89	0.7649
	Transacting protein	WGVVCILPW	HLA-B*53:01	81	89	0.47
LMWGVVCIL		HLA-B*14:02	79	87	0.85	1.5548
VIAPQPRPL		HLA-B*15:09	68	76	0.80	1.2564
IAPQPRPLL		HLA-C*03:03	69	77	0.34	1.2724
RRKWPHLKP		HLA-B*73:01	46	54	0.22	1.0980
PQPRLLLL		HLA-B*40:13	71	79	0.73	0.9919
AHPAISAEQ		HLA-B*15:09	56	64	0.86	0.9307
APQPRLLL		HLA-B*08:02	70	78	0.28	0.8880
HPAISAEQL		HLA-B*15:09	57	65	0.09	0.8577
RKWPHLKPL		HLA-B*27:05	47	55	0.54	0.6528
PLLLMWGV		HLA-A*02:12	75	83	0.30	0.6014
QLLAVIAPQ		HLA-A*02:16	64	72	0.69	0.5690
GTRVWPLGR		HLA-A*31:01	114	122	0.42	0.5094
ISAEQLLAV		HLA-A*69:01	60	68	0.35	0.4689
PLAHPAISA		HLA-A*02:50	54	62	0.85	0.4175

**Table 4.** The predicted T cytotoxic cells epitopes, their antigenicity scores from the polymerase, envelope, and transacting factor proteins. \*PR: Percentile rank with a score of  $\leq 1$ . <sup>#</sup>The Vaxijen server for antigenicity threshold was 0.4. All the predicted epitopes were nonallergic and nontoxic.

### Solubility of the assembled vaccine

Based on the Protein-Sol server, a scaled solubility score of 0.499 was obtained for the vaccine construct, compared with 0.45 for the population solubility of *E. coli* (Fig. 4a). As a confirmation, SOLpro was further used to predict the solubility. The probability of the proposed vaccine upon expression on SOLpro was 0.9843, greater than 0.5, provided by the server.

### Stability of the assembled vaccine

By engineering disulfide bonds into the structure of the proposed vaccine, the structural stability of the vaccine was improved. The improvement in stability was made possible by substituting the amino acids in the highly mobile regions in the sequence of the vaccine by cysteine residues. As per the Disulfide by Design 2.0 server, 94 amino acid pairs were identified to form disulfide bonds. However based on the Chi3 angle between +97 and -87 and a tolerance of 30 and a maximum Ca-Cb-S angle of 114.60 in the server, five pairs of residues (amino acids) were unstable regions and were replaced by cysteine-cysteine residues. The position and the replaced residues in the vaccine structure were A107-R127; I150-G179; P210-P280; P278-P312 and G500-L538 and were shown in Fig. 4b,c.



Protein	Core sequence	Alleles	Start	End	Peptide sequence	*PR	#Vaxijen antigenicity
Polymerase	VLDLKDCCF	HLA-DPA1*01:03/DPB1*04:02	99	116	RGWPLMVLDLKDCCFFSIP	0.42	1.4679
	LEPLRLKHP	HLA-DRB1*12:01/HLA-DRB1*11:01	161	178	VGQVLEPLRLKHPSLRML	2.60	1.1931
	PVFVIRKAS	HLA-DQA1*05:01/DQB1*04:02	53	70	CWNTPVFVIRKASGSYRL	2.50	1.0881
	LTLITKLR	HLA-DRB4*01:03	364	381	VLTLITKLRASAVRTFG	0.61	1.0175
	LLITKLRAS	HLA-DRB1*08:01	365	382	LTLITKLRASAVRTFGK	1.30	0.8702
	APVGLVAEP	HLA-DQA1*05:01/DQB1*03:02	230	247	LGSTYVAPVGLVAEPRIA	3.10	0.887
	PLPEGKLV	HLA-DPA1*01:03/DPB1*04:02	17	34	TPVWIDQWPLPEGKLVAL	2.90	0.8656
	KDCFFSIPL	HLA-DQA1*01:01/DQB1*05:01	102	119	PLMVLDLKDCCFFSIPLAE	0.71	0.7918
	ARAVAMALL	HLA-DPA1*01:03/DPB1*03:01	481	498	QQLEARAVAMALLLWPTT	1.20	0.7845
	QQGAPVLSA	HLA-DQA1*01:02/DQB1*06:02	80	97	KLVPGAVQQGAPVLSAL	2.50	0.776
Envelope	AIDFLLAH	HLA-DPA1*02:01/DPB1*01:01	470	487	RHAVLQNRAAIDFLLAH	3.00	1.0368
	TVVTADRHN	HLA-DPA1*01:03/DPB1*04:02	228	245	FTVVTADRHNLFMGSEYC	2.90	1.0245
	QNRAAIDFL	HLA-DQA1*01:02/DQB1*05:02	472	489	AVLQNRAAIDFLLAHGH	2.90	0.9233
	YLGKLTMLA	HLA-DRB1*10:01	368	385	GPCYLKGLTMLAPKHTDI	1.70	0.8309
	ELQLGSQS	HLA-DRB4*01:01	165	182	PHLQLLGSQSLPNITNI	3.20	0.6819
	VLQNRAAID	HLA-DRB1*13:02	468	485	SIRHAVLQNRAAIDFLLL	1.90	0.4838
Transacting	WGVVILPW	HLA-DRB1*04:02	75	92	PLLLLMWGVVILPWRGW	2.60	1.5006
	VIAPQPRPL	HLA-DRB3*03:01	63	80	EQLLAVIAPQPRPLLLM	2.70	1.2564
	AHPAISAEQ	HLA-DQA1*05:01/DQB1*03:02	49	66	WPHLKPLAHPAISAEQLL	2.20	0.9307
	RKWPHLKPL	HLA-DPA1*02:01/DPB1*14:01	41	58	MRRRWRKRWPHLKPLAHP	2.10	0.6528
	LLAVIAPQP	HLA-DRB1*10:01	60	77	ISAEQLLAVIAPQPRPL	2.60	0.5912
	LLLLMWGVV	HLA-DRB1*04:03	76	93	LLLLMWGVVILPWRGWE	2.10	0.5785
	PLAHPAISA	HLA-DQA1*05:01/DQB1*03:03	49	66	WPHLKPLAHPAISAEQLL	1.90	0.4175

**Table 5.** The predicted T helper cells epitopes and their antigenicity scores from the polymerase, envelope, and transacting factor proteins. \*PR: Percentile rank with a score of  $\leq 10$ . #The Vaxijen server for antigenicity threshold was 0.4. All the predicted epitopes were nonallergic and nontoxic.

### Immune simulation

The obtained immune simulation results were coincided with actual immune responses. This was proved by marked increase in the primary, secondary and tertiary immune responses accompanied by drop in the antigen concentration (Fig. 5a). The cytokines and interleukins (IL) levels during the injections showed that the IL-2 level was compatible with the measure of diversity (Simpson index, D) (Fig. 5b). The elevation of the measure of diversity over time is considered as danger signal together with leukocyte growth factor. Therefore, the lower the measure of diversity value, the lower the diversity. In addition the primary response, for instance, was featured by augmented IgM level, while, secondary and tertiary responses provided marked elevation in the population of B-cells and the antibodies level (Fig. 5c). This showed the development of immune memory accompanied by rapid clearance of the antigen upon subsequent exposures. Moreover the population of T-cytotoxic (TC) (Fig. 5d) and T-helper (TH) (Fig. 5e) lymphocytes showed high response level coincided with memory development. The natural killer cells maintained high levels throughout the duration of exposure (Fig. 5f).

### Molecular dynamic simulation (MD)

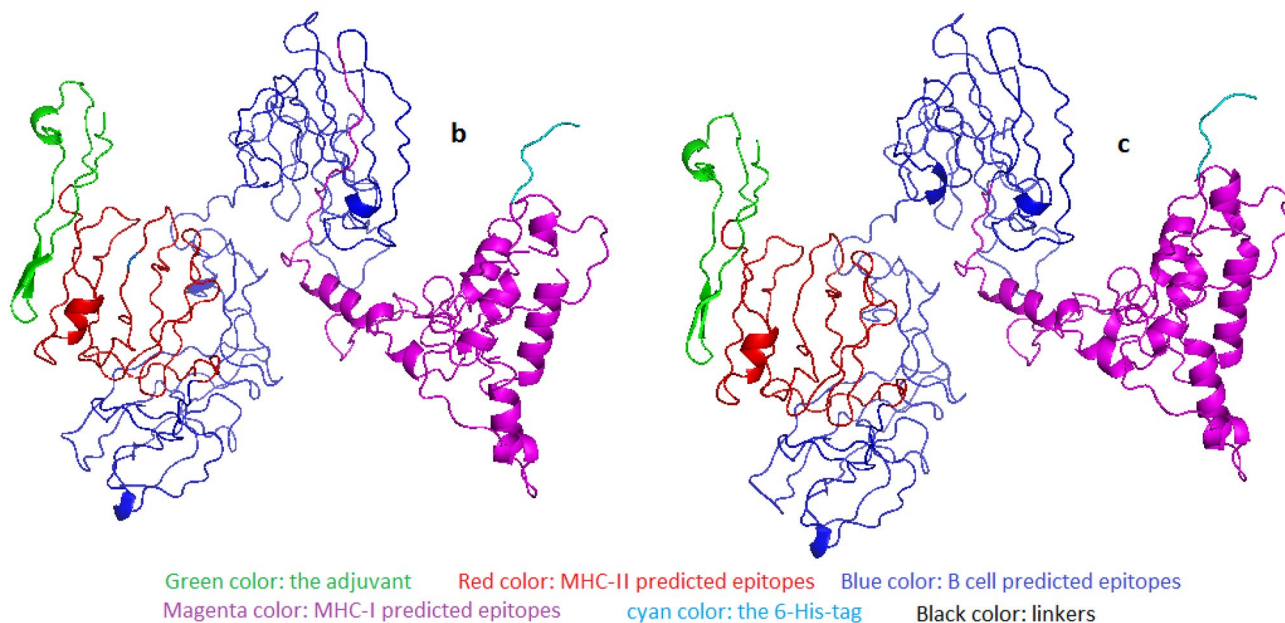
A Normal mode analysis (NMA) was performed on the MD of the vaccine protein using the iMODS server. As shown in Fig. 6a, the arrows indicated the direction in which each vaccine protein residue moves. Deformability was also demonstrated with hinges in the chief chain, as a result of an individual distortion of the residues (Fig. 6b). Experimental B-factors were calculated on the basis of the PDB field and the NMA data (Fig. 6c). A normal mode of deformability of the vaccine structure was shown by the eigenvalue, which directly correlated to the energy required with the deformability. The obtained eigenvalue ( $7.182836e-07$ ) demonstrated the stiffness of the motion (Fig. 6d). The lower eigenvalue is always associated with the easier deformation of the protein structure. The normal mode variance is inversely related to the eigenvalue. Figure 6e illustrated the cumulative variance and individual variance as green and purple bars, respectively. It was possible to determine the correlations between proteins by examining the covariance matrix (Fig. 6f). Thus, red identified correlated motions, white indicated uncorrelated motions, and blue indicated anti-correlated motions. Spring-connected or joined atom pairs were demonstrated in the elastic network model. A single-atom pair spring was represented as a dot, and colored according to its stiffness, with darker dots denoting stiffer strings, and vice versa (Fig. 6g).

### Discontinuous B-cell epitopes prediction

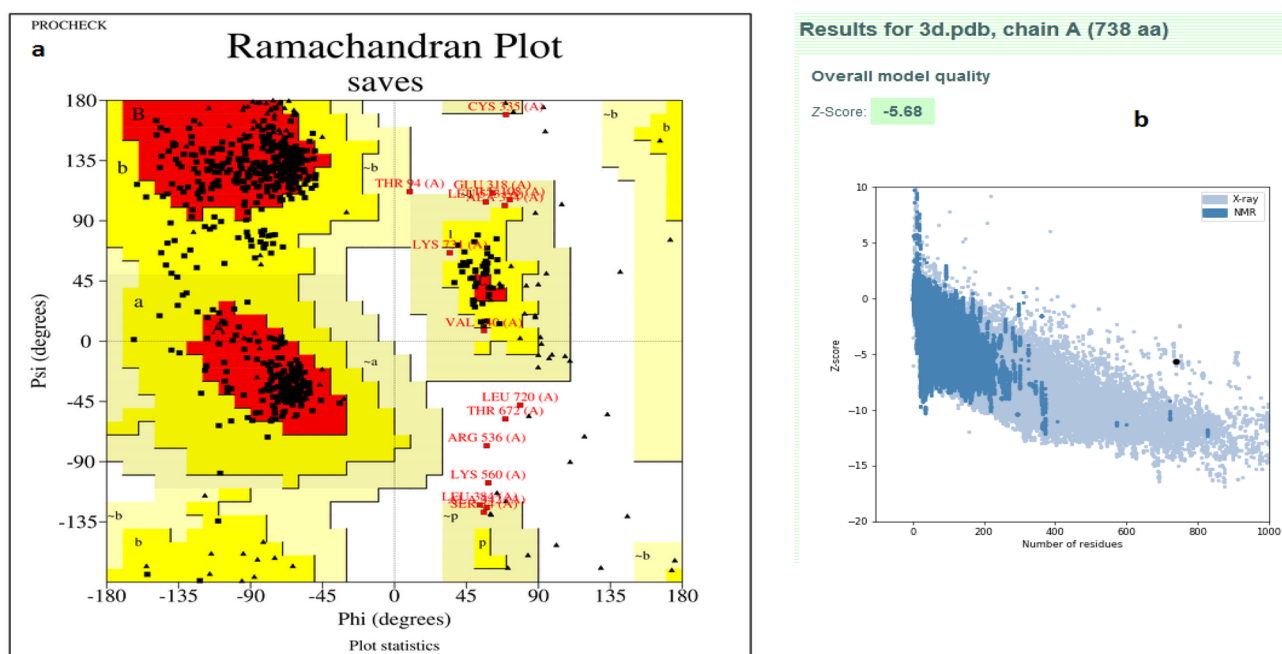
Table 6 and Fig. 7 demonstrated six discontinuous B cell epitopes. The scores of these epitopes were ranged from 0.996 to 0.615 with a total of 405 predicted residues. The size of the conformational epitopes ranged from 4 to 108 residues.

**The final vaccine construct**

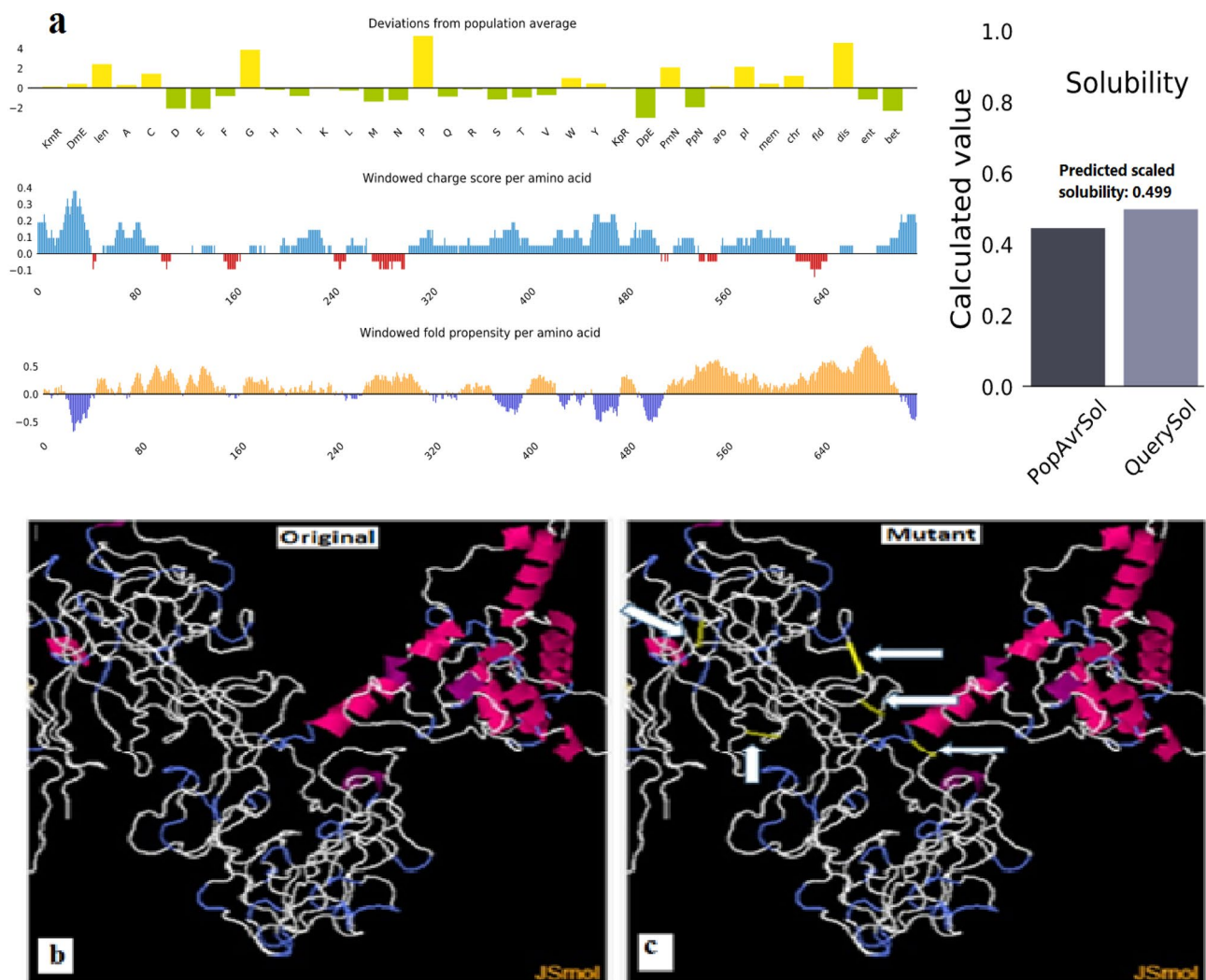
5-GIINTLQKYICRVRRGRCVAVLSCLPKKEEQIGKCKSTRGRKCCRKKEEAAK VLDLKDCCFFGPGPGLLEPLRLKHPGPGPGPVFVIRK  
 ASGPGPGLTLLITLKRGPGGGAIDFLLAHGPGPGTVVADRHNPGPGWGVCILPWGPGPGVIAPQPRPLGPGPGVEKELQLG  
 HIEPGGPGVNAKLVFPGAVQGGPGWNLDMKMAWREIGPGPGAGVNPRLGLPLQGGPGIARPLHVS LKVRGPGPGIWQT  
 DFTLEPRMGPGGTLEPRMAPRSWLGPGPLMVLDLKDCFFGPGGLDLKDCFFSIPLGPGPGVLEPLRLKHPSLGPGPGKACNI  
 SMQARERGGPGVYGRFPQGVVWYLGPGPGACWSVKQANLTGPGPGLTGYPGKTSKKGDPGPGKTPLLPRTVNYIGPGPGISGI  
 TGGCVGFRGPGPGVNCQVGRQYRCGPGPLWITWANRTGQTGPGPGTGIARRKSTSHLGPGLPTRVNYIIGGGPGPGKGT  
 SKKDSKEKGPGGPGHIEPSLSCYAAHIGPRALSKEYAALEARAVAMA YAAVLDLKDCCFFYAAAMA AVLHVR YAARCEQGAIGVYA  
 ARLACWSVKQYAAACWSVKQANLYAAWSVKQANLTAA LQNRAAIDFYAADFCLSTQSA YAAAIDFLLA YAA SVKQANLTY  
 AA WGVVILPWYAALMWGVVILYAAVIAPQPRPL YAAIAPQPRPLLYAARRKWP H LKPHHHHHH -3



**Figure 2.** (a) The primary sequence of the proposed vaccine. (b) The tertiary structure of the vaccine predicted by Raptor X server. c The refined structure of the vaccine predicted by the Galaxy web server.



**Figure 3.** (a) In the Ramchandan plot the most favoured region comprised 90.9%; additional allowed region comprised 6.1%, generously allowed region comprised 1.9%, a disallowed region comprised 1.0% of the residues. (b) ProSA-server with Z-score of -5.68.



**Figure 4.** (a) The vaccine solubility in comparison to the solubility of *E. coli*. (b) Stability of the vaccine protein before disulfide bond engineering in the original form (the form before substitution of amino acids by cysteine). (c) The mutant form (the form after substitution of amino acids by cysteine) with five pairs of disulfide bond formation. The disulfide pairs were shown by golden sticks and pointed by white arrows.

#### Active site detection in the vaccine structure

Prior to docking the active binding pocket of the vaccine construct was predicted by CASTp. The highest pocket area and volume were 7086.109 Å<sup>2</sup> and 18,104.055 Å<sup>3</sup>, respectively. This result indicated pocket area and volume were suitable for binding during the molecular docking. The surface binding pocket of the vaccine and the sequence were shown in Fig. 8.

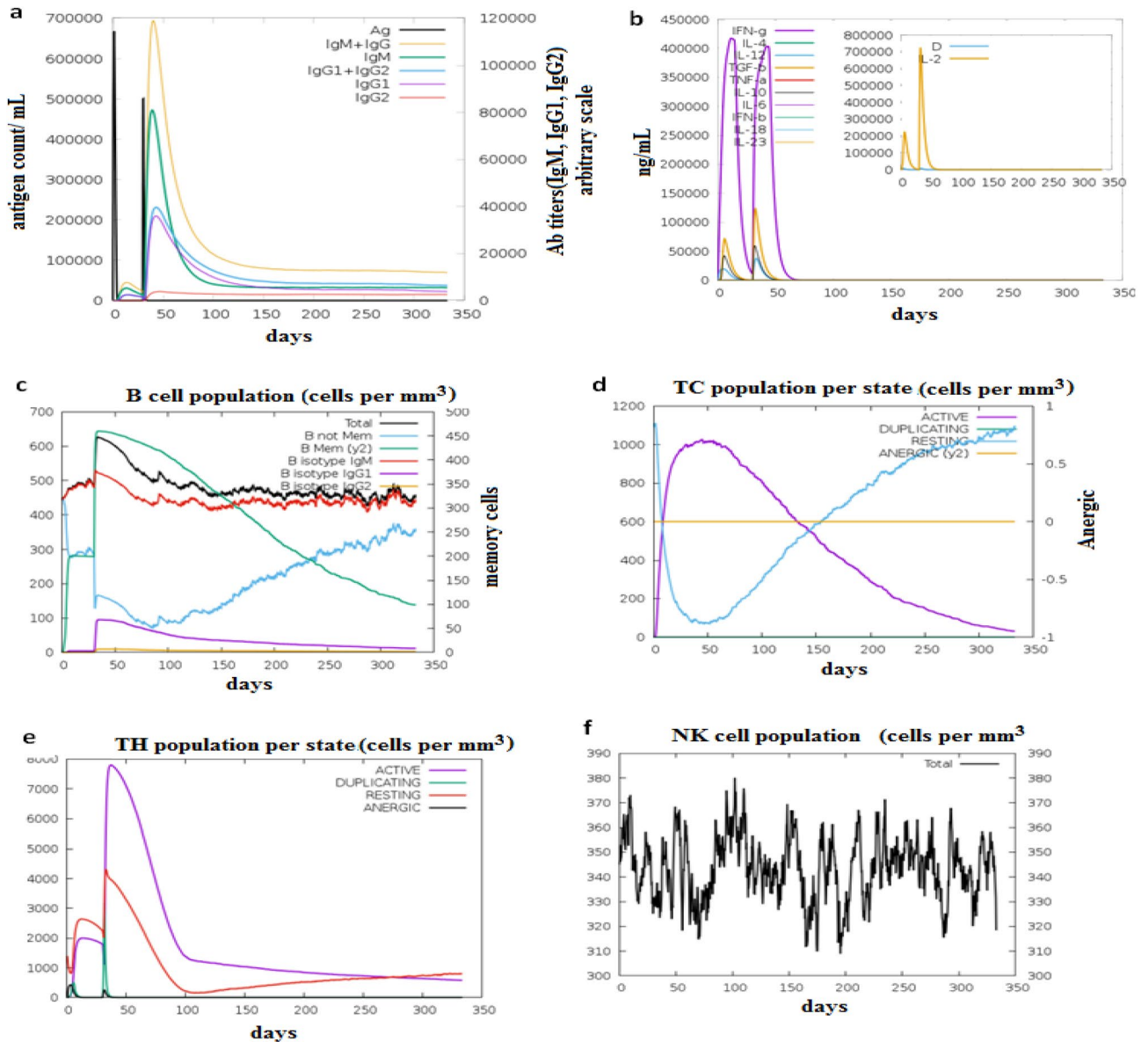
#### Molecular docking of the vaccine protein with chicken TLR7

The interaction between the vaccine construct and chicken TLR7 was assessed by HADDOCK software. HADDOCK clustered 13 structures in 3 cluster(s), which represents 6% of the water-refined models. Upon refinement, 20 structures were grouped into one cluster, resulting in 100% of the HADDOCK water-refined version. The binding affinity between the vaccine and the chicken TLR7 was  $-263.0 \pm 3.1$  demonstrating the strong binding between the molecules. As shown in Fig. 9, this binding was evident by 20 hydrogen bonds, 2 salt bridge, and 184 non-bonded contacts. These bonding events between the amino acids of the molecules were provided in Table 7. Additionally PRODIGY web server showed binding affinity in terms of Gibbs free energy ( $\Delta G$ ) and thermodynamics (dissociation constant) between the docked molecules. Such kind of binding affinity decided the real interaction between the docked molecules under certain circumstances within the cell. The server showed  $\Delta G$  values  $-21.1$  kcal/mol for the vaccine construct and chicken TLR7 and the dissociation constant was  $3.1 \times 10^{-16}$  indicating the docked molecules were energetically viable.

#### In silico molecular cloning

The potential host expression of the target protein was performed by JCAT. The protein sequence of the vaccine was reversibly translated into DNA sequence. The index codon adaptation value of the DNA sequence was equal



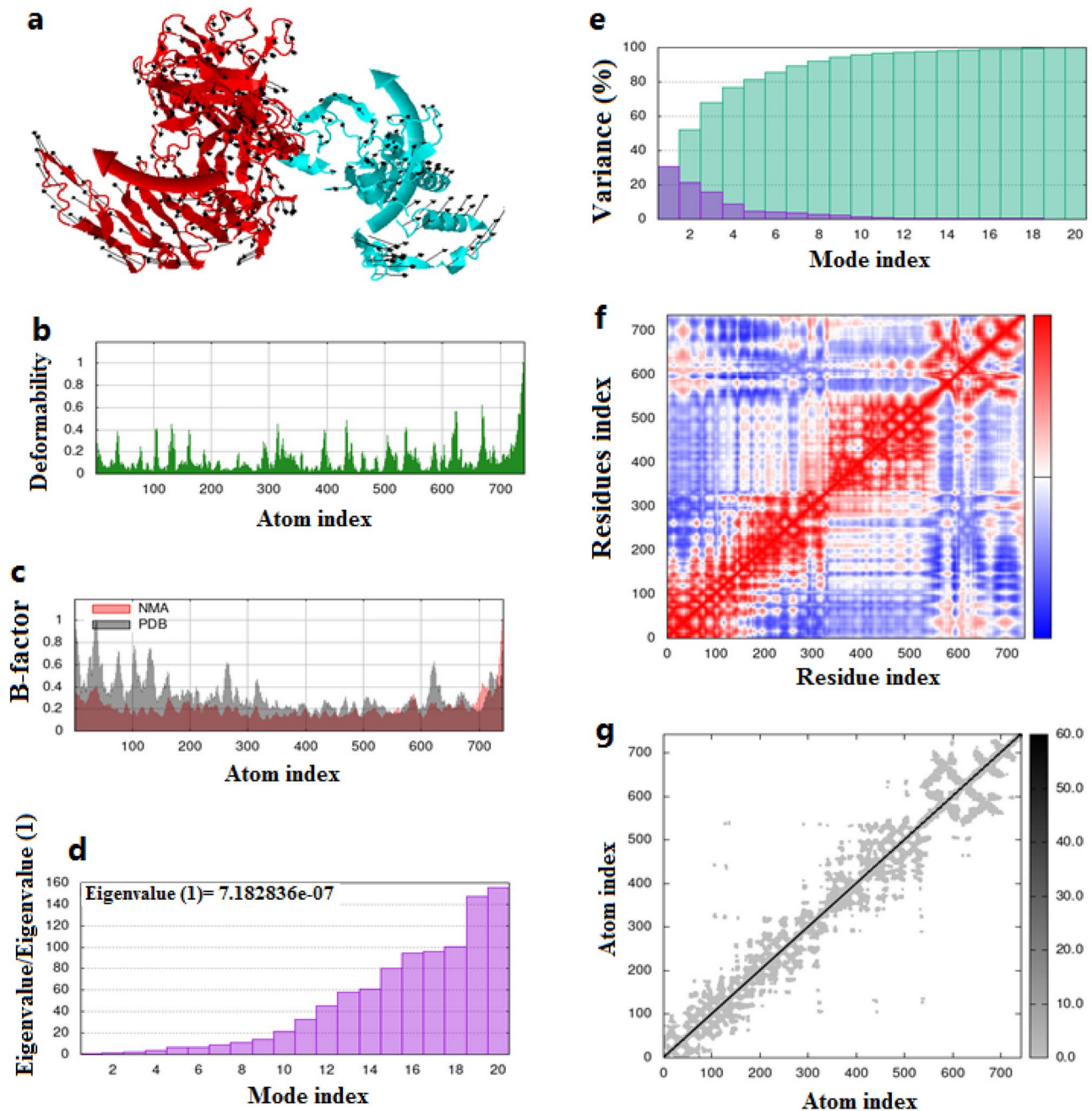


**Figure 5.** The immune simulation of the predicted vaccine after the two injections of the antigen. (a) Antibodies production in response to antigen injections (antibodies were shown as different colored peaks and the antigen was shown in black color). (b) The induced cytokines secretion and the IL-2 level with the measure of diversity. (c) Showed the memory, not memory and the isotypes of B-cell populations. (d) Showed the active T-cytotoxic (TC) cell populations. (e) Showed the active T-helper (TH) cell populations. In (d, e) The resting state demonstrated the cells not provided with the antigen (vaccine). The anergic state demonstrated tolerance of the T-cells to the antigen due to repeated exposures. (f) Natural killer cell populations.

to 1.0, indicating a higher proportional abundance of the common codons. The improved GC content value of the DNA sequence was 59.53%, providing significant GC content. As shown in Fig. 10 the DNA sequence was cloned into the pET-30a(+) vector utilizing Snapgene software between the restriction enzyme cutting locations of BamH1 and Xho1.

## Discussion

The most common avian retrovirus that causes a variety of neoplastic diseases in chicken is the avian leukosis virus (ALV)<sup>2</sup>. Globally, the ALV morbidity and mortality rates contributed to the poultry industry's economic decline<sup>1,3</sup>. This is accompanied by subsequent adverse effects on the food supply worldwide. Preventing and controlling viral infection in avian industry is always via mass vaccination means. Therefore vaccines designed to combat avian viral diseases will significantly alleviate selection pressure on the virus and on the field strains<sup>68</sup>. Concerning ALV infection in poultry, many anti-ALV vaccines were developed, but they targeted only specific strains. Also some of the vaccine trials had less immunogenicity and limited protection<sup>69</sup>. Currently, neither known treatment nor vaccination against ALV is available. Multiple studies used the multi-epitope vaccine prediction against ALV and evaluated their possibility as effective vaccine candidate via challenging in chickens<sup>70-72</sup>.



**Figure 6.** Showed the MD of the vaccine protein complex. (a) The direction of the motion was shown by the red and cyan colors. (b) The stability of the vaccine was analyzed by the low main chain deformability. (c) The B factor/ mobility. (d) The Eigenvalue demonstrated the protein's normal mode and the stiffness of the motion. (e) The normal mode variance and (f) is the covariance matrix. (g) the elastic network model showed a stiffer mode of the residues.

For instance, one study provided a novel oral vaccine of recombinant gp85 protein in *L. plantarum* with a significant increase in antibodies post-inoculation<sup>72</sup>. The study demonstrated a protection against ALV-J and showed protective immune response against early ALV-J infection based on viremia analysis<sup>73</sup>. Another study showed the impact of polysaccharides from *Ulvapertusa* as anti-ALV-J. The polysaccharides demonstrated strongest suppression of the ALV-J activity as they bound with the viral particles and obstacle ALV-J adsorption by host cells accompanied by significant reduction of gp85 protein expression<sup>74</sup>. However these studies reported partial immune protections against ALV-J infections in chickens.

In this study a vaccine with multi-epitopes was designed and showed increased immunogenicity and enhanced immune responses as a result of the existence of epitopes from various target genes. Also the designed vaccine activated the humoral and cell-mediated immunity as previously described<sup>75</sup>. These could solve limitations occurred during controlling the ALV infection<sup>75,76</sup>. Most importantly, the safety and effectiveness, allergenicity

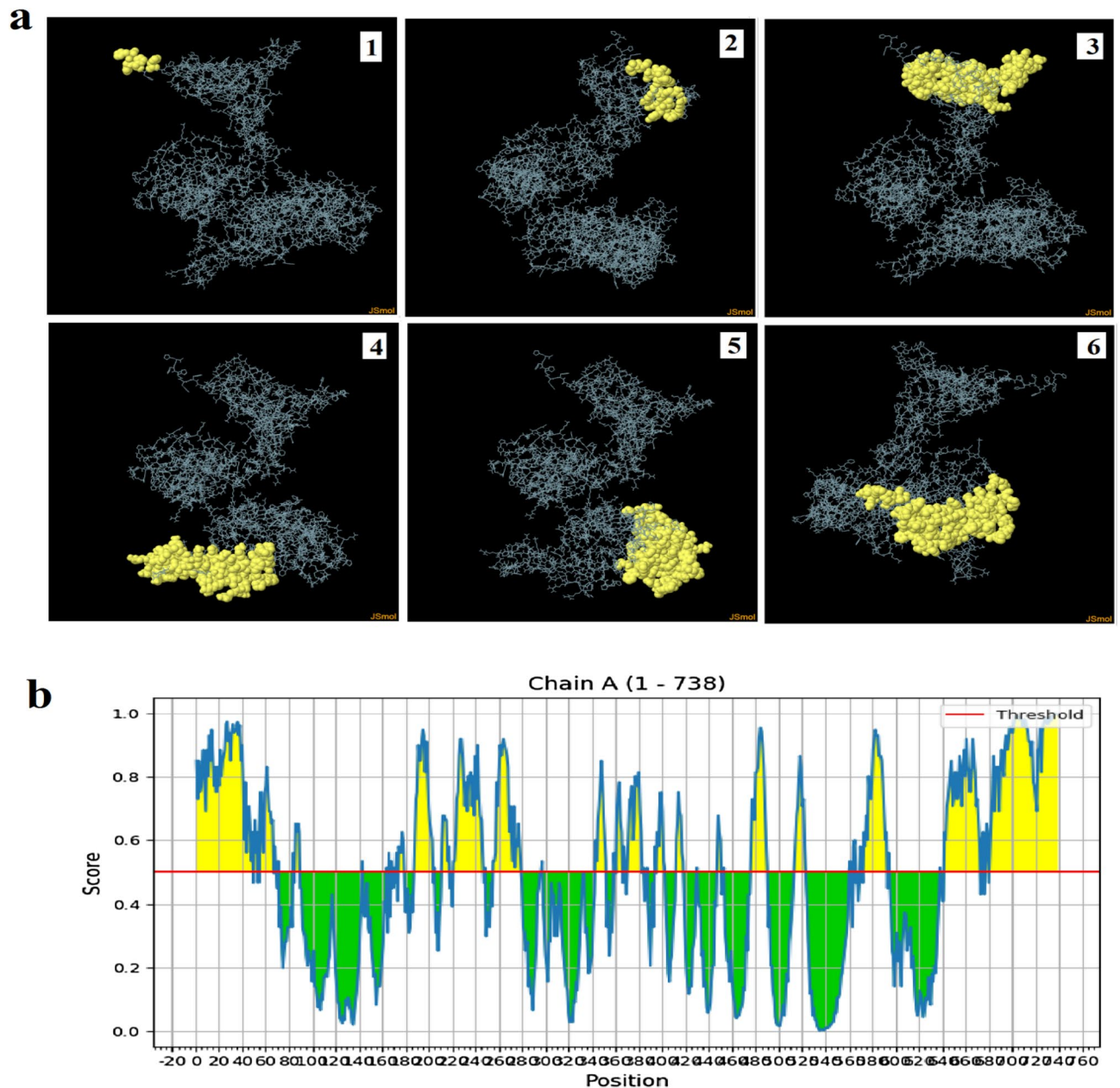
No	Residues	Number of residues	Score
1	A:H734, A:H736, A:H737, A:H738	4	0.996
2	A:R717, A:P718, A:L719, A:L720, A:Y721, A:A722, A:A723, A:R724,	16	0.891
3	A:K560, A:D561, A:F564, A:Y565, A:A566, A:A570, A:A571, A:L573, A:H574, A:V575, A:R576, A:Y577, A:A578, A:A579, A:R580, A:C581, A:E582, A:Q583, A:G584, A:A585, A:I586, A:G587, A:V588, A:Y589, A:A590, A:A591, A:R592, A:L593, A:A594, A:C595, A:C642, A:L643, A:S644, A:T645, A:Q646, A:S647, A:A648, A:Y649, A:A650, A:A651, A:A652, A:A653, A:I654, A:D655, A:F656, A:L657, A:L658, A:L659, A:A660, A:Y661, A:A662, A:A663, A:S664, A:V665, A:K666, A:Q667, A:A668, A:N669, A:L670, A:T671, A:T672, A:A674, A:A675, A:G677, A:V678, A:V679, A:I681, A:L682, A:P683, A:W684, A:Y685, A:A686, A:A687, A:L688, A:M689, A:W690, A:G691, A:V692, A:V693, A:C694, A:I695, A:L696, A:Y697, A:A698, A:A699, A:V700, A:I701, A:A702, A:P703, A:Q704, A:P705, A:R706, A:P707, A:L708, A:Y709, A:A710, A:A711, A:I712, A:A713, A:P714, A:Q715, A:P716	102	0.78
4	A:G1, A:I2, A:I3, A:N4, A:T5, A:L6, A:Q7, A:K8, A:Y9, A:C11, A:R12, A:V13, A:R14, A:G15, A:G16, A:R17, A:C18, A:A19, A:V20, A:L21, A:S22, A:C23, A:L24, A:P25, A:E28, A:Q29, A:I30, A:G31, A:K32, A:C33, A:S34, A:T35, A:R36, A:G37, A:R38, A:K39, A:C40, A:C41, A:R43, A:E46, A:A47, A:A48, A:A49, A:K50, A:V51, A:L52, A:D53, A:K55, A:D56, A:C57, A:F58, A:F59, A:G60, A:P61, A:G62, A:P63, A:G64, A:L65, A:E66, A:P67, A:L68, A:L70, A:K71, A:H72, A:V82, A:R84, A:K85, A:A86, A:S87, A:G88, A:P89, A:G90, A:P91	73	0.738
5	A:W143, A:G144, A:L167, A:Q168, A:L169, A:G170, A:H171, A:I172, A:E173, A:P174, A:G175, A:P176, A:G177, A:P178, A:G179, A:N181, A:A182, A:G188, A:A189, A:V190, A:Q191, A:G192, A:P193, A:G194, A:P195, A:G196, A:W197, A:N198, A:L199, A:D200, A:M201, A:K202, A:M203, A:A204, A:W205, A:E207, A:P210, A:G211, A:P212, A:G213, A:A214, A:G215, A:V216, A:N217, A:P218, A:R219, A:G220, A:L221, A:G222, A:P223, A:L224, A:Q225, A:G226, A:P227, A:G228, A:P229, A:G230, A:I231, A:A232, A:R233, A:P234, A:L235, A:H236, A:V237, A:S238, A:L239, A:K240, A:V241, A:R242, A:G243, A:P244, A:G245, A:P246, A:G247, A:I248, A:Q250, A:F253, A:T254, A:L255, A:E256, A:P257, A:R258, A:M259, A:G260, A:P261, A:G262, A:P263, A:G264, A:T265, A:L266, A:E267, A:P268, A:R269, A:M270, A:A271, A:P272, A:R273, A:S274, A:W275, A:L276, A:G277, A:P278, A:G279, A:P280, A:G281, A:P309, A:P312, A:G313	108	0.647
6	A:A342, A:E344, A:G345, A:P346, A:G347, A:P348, A:G349, A:V350, A:G351, A:F352, A:G356, A:P358, A:W359, A:Y360, A:L361, A:G362, A:P363, A:G364, A:P365, A:G366, A:A367, A:C368, A:W369, A:S370, A:V371, A:K372, A:Q373, A:A374, A:N375, A:L376, A:T377, A:T378, A:G379, A:P380, A:G381, A:P382, A:G383, A:L384, A:T385, A:G386, A:Y387, A:P388, A:K394, A:D395, A:G396, A:P397, A:G398, A:P399, A:G400, A:K401, A:T402, A:P403, A:V409, A:N410, A:Y411, A:I412, A:G413, A:P414, A:G415, A:P416, A:G417, A:I418, A:S419, A:R429, A:G430, A:P431, A:R445, A:C446, A:G447, A:P448, A:G449, A:P450, A:G451, A:L452, A:W453, A:R475, A:S476, A:T477, A:S478, A:H479, A:L480, A:G481, A:P482, A:G483, A:P484, A:G485, A:L486, A:P487, A:T488, A:R489, A:V490, A:N491, A:Y492, A:K514, A:G515, A:P516, A:G517, A:P518, A:G519, A:G520, A:H521, A:I522	102	0.615

**Table 6.** The number of the predicted discontinuous B cell epitopes with the number of the residues and their scores.

and the immunogenicity of the predicted vaccine were also taken into consideration to ensure the safety of the designed epitopes<sup>77</sup>. In addition, the toxic effect, the solvent accessibility of the amino acids, the identification of B cells, and MHC molecules were also contemplated to ensure the effectiveness of the predicted epitope vaccine. All these measures give the predicted vaccine an advantage over the traditional ones for controlling the ALV infection.

Thus the conserved predicted epitopes from ALV proteins were submitted to the ABCpred server. Based on ANN, Hidden Markov model (HMM) and support vector machine (SVM) in the ABCpred server the B cell epitopes were predicted<sup>78</sup>. Furthermore, the predicted epitopes were subjected to antigenic, allergenic and toxic analysis to confirm their suitability as B cell epitopes. Also T cell epitopes were predicted from their reference sequences using the IEDB server. In addition to their high binding affinity to MHC alleles, the predicted epitopes demonstrated high antigenicity score in VaxiJen server, and they revealed no allergic or toxic characteristics. Therefore they were picked to enter the vaccine protein structure. With the aid of expedient linker sequences (protein spacers), the generated B- and T-cell epitopes were fused together<sup>49,52</sup>. Linkers are crucial to the assembly of stable, bioactive fusion proteins. Essentially, linkers reduced the likelihood of junctional antigen formation as well as enhancing antigen processing and presentation<sup>52</sup>. They are also important to construct and facilitate structural flexibility and reduced rigidity<sup>52</sup>. A sequence with the least junctional immunogenicity was generated in this study using the linkers GPGPG and YAA. The GPGPG linkers were applied to facilitate immune processing and merge the B-cells and T-helper cell epitopes. The YAA linkers ameliorated the immunogenicity of a vaccine by impacting protein stability and epitope presentation capacity and were used to fuse the cytotoxic T-cells<sup>24,79</sup>. As an adjuvant, the  $\beta$ -defensin was added via an EAAAK linker at the N terminus of the vaccine construct to improve the immunogenicity of the vaccine. EAAAK are helical linkers used to control the distance



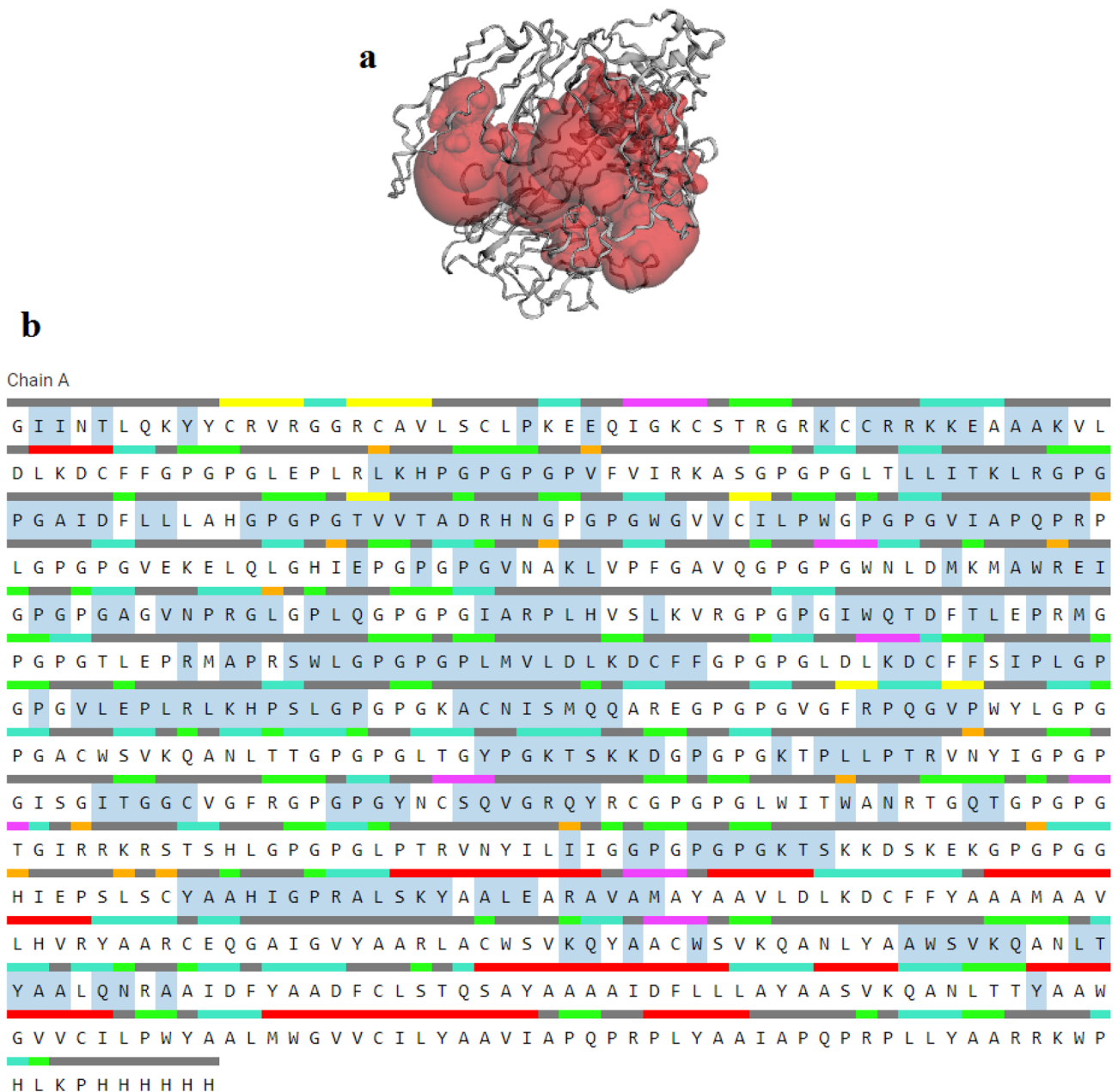


**Figure 7.** (a) showed the 3D structures of six discontinuous B-cell epitopes predicted by the ElliPro (1–6). Epitopes were shown in yellow color, while grey color showed the constructed vaccine. (b) The yellow color demonstrated the discontinuous epitopes while the green color was the continuous epitopes. The red line showed the threshold of the residues score.

and decrease the interference between the domains<sup>24,79</sup>. As a 45amino acids peptide with a relatively small size, the  $\beta$ -defensin was used for its immune modulation and antimicrobial features<sup>44</sup>. To facilitate purification and downstream testing, a small 6His tag was added to the proposed vaccine at the C-terminal to prevent protein structure from being altered<sup>80</sup>.

The stability of the vaccine was confirmed by the ProtParam server based on its physiochemical properties. VaxiJen and AllerTOP servers were used to assess the antigenic and allergenic features of the vaccine. The results indicated that the vaccine was antigenic without causing any allergic reactions. In order to select the best score of the model generated by the 3D structure of the vaccine protein, the secondary and tertiary structures of the vaccine construct were analyzed. The Ramachandran plot showed favorable results in the distribution of the vaccine residues and provided a stable structure. The ProSA server indicating that the overall model is suitable for acceptance as a potential ALV vaccine<sup>51,52</sup>.

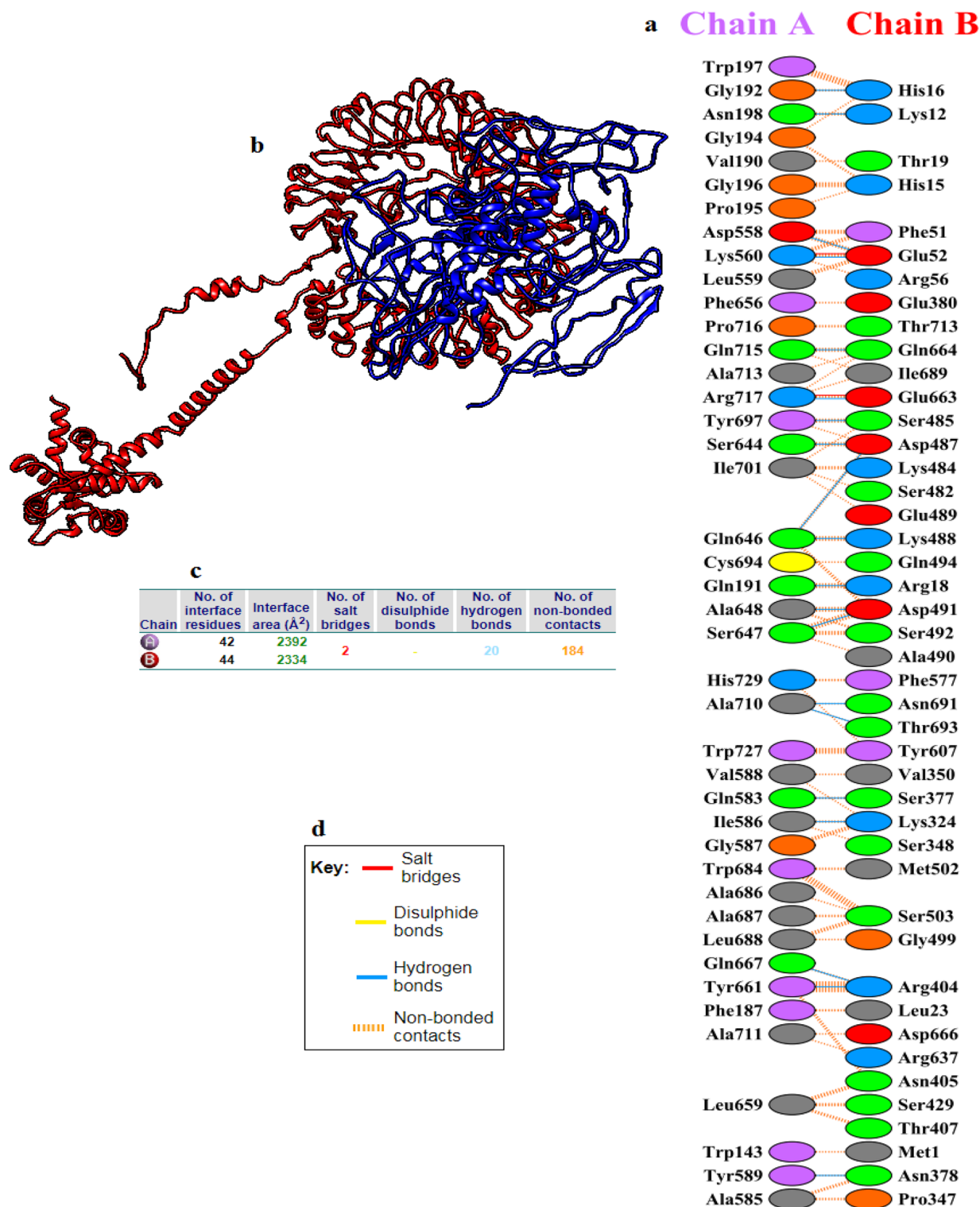
The solubility of the designed vaccine in this study was calculated with the protein-sol and SOLpro servers. As a comparison with the solubility of *E. coli*, Protein-sol presented the vaccine as a soluble protein and predicted a scaled solubility of 0.499, an increase over 0.45 from the average solubility of the *E. coli* population. According to the SOLpro server, the predicted solubility was 0.9843, which confirmed this result. To obtain disulfide bonds



**Figure 8.** (a) The pocket panel (shown in red color) in the structure of the vaccine. (b) the sequence and annotation panels in the vaccine construct.

between the vaccine residues, the proximity and geometry composition of the residue pairs were evaluated for the formation of disulfide bonds. Five unstable regions in the vaccine structure were replaced by the formation of disulfide bonds. Disulfide bonding increases the stability of the vaccine protein as previously stated<sup>51,52</sup>.

Immune simulation demonstrated results that consistent with the real immune responses. Generally there were elevated levels of the immune responses after repeated exposure to the vaccine (antigen). In addition, there was marked development in the memory cells of B and T lymphocytes. Most importantly, IL-2 and IFN- $\gamma$  were elevated following the initial injection and provided peak levels after antigen repeated exposures, showing the high levels of T-helper lymphocytes and efficient immunoglobulin production. The Simpson index demonstrated a possible different immune response, indicating the vaccine structure contains multiple B and T cells epitopes<sup>44</sup>. A study by Landman et al., demonstrated the interaction of the NK cells during ALV infection<sup>81</sup>. They showed that during ALV infection in immunosuppressed chicken, the NK cells provided reduced killing activity than the NK cells of the uninfected controls. Natural killer cells play a paramount defense mechanism in host and surveillance of tumor, resulting in cell death and secretion of cytokines and chemokine. Moreover, NK cells have a significant role in immune regulation of T cells and DC functions during viral infection in mouse models<sup>81</sup>. In addition to that, there is scarcity in ALV vaccine researches concerning the immune system of chickens. Thus



**Figure 9.** Molecular docking interaction between the vaccine construct with chicken TLR7. (a) Interacting residues between the vaccine (chain A) and TLR7 (chain B). (b) Chicken TLR7 (red color) and the vaccine construct (blue color) docked complex. (c) interface statistics result. (d) Key showing the residue interactions across interface between the docked molecules.

vaccine research studies may provide a realistic importance and increase the knowledge of immune system of ALV with great insights into human retroviral diseases<sup>69</sup>.

			Vaccine			Hydrogen bonds			Chicken TLR 7			
	Atom number	Atom name	Residue name	Residue number	Chain	<->	Atom number	Atom name	Residue name	Residue number	Chain	Distance
1	1687	OE1	GLN	191	A	<->	6785	NH1	ARG	18	B	2.73
2	1697	O	GLY	192	A	<->	6766	NE2	HIS	16	B	3.2
3	1743	OD1	ASN	198	A	<->	6723	NZ	LYS	12	B	2.7
4	4893	OD2	ASP	558	A	<->	7105	N	GLU	52	B	2.71
5	4912	NZ	LYS	560	A	<->	7112	OE2	GLU	52	B	2.59
6	5133	OE1	GLN	583	A	<->	10,298	OG	SER	377	B	3
7	5158	O	ILE	586	A	<->	9786	NZ	LYS	325	B	2.76
8	5182	OH	TYR	589	A	<->	10,308	ND2	ASN	378	B	3.05
9	5715	OG	SER	644	A	<->	11,439	OD1	ASP	487	B	2.62
10	5734	OE1	GLN	646	A	<->	11,450	NZ	LYS	488	B	2.69
11	5735	NE2	GLN	646	A	<->	11,440	OD2	ASP	487	B	3.02
12	5740	N	SER	647	A	<->	11,477	OD1	ASP	491	B	2.85
13	5748	N	ALA	648	A	<->	11,477	OD1	ASP	491	B	2.7
14	5865	OH	TYR	661	A	<->	10,580	NH1	ARG	404	B	2.98
15	5916	OE1	GLN	667	A	<->	10,580	NH1	ARG	404	B	2.79
16	6195	OH	TYR	697	A	<->	11,417	OG	SER	485	B	2.72
17	6312	O	ALA	710	A	<->	13,440	ND2	ASN	691	B	3
18	6312	O	ALA	710	A	<->	13,458	OG1	THR	693	B	3.11
19	6348	NE2	GLN	715	A	<->	13,192	OE1	GLN	664	B	2.83
20	6369	NH1	ARG	717	A	<->	13,182	OE21	GLU	663	B	3.01
			Vaccine			Salt bridges			Chicken TLR 7			
1	4912	NZ	LYS	560	A	<->	7111	OE1	GLU	52	B	2.59
2	6369	NH1	ARG	717	A	<->	13,183	OE2	GLU	663	B	3.01

**Table 7.** List of Atom – Atom Interactions between the Vaccine and chicken TLR7 Interface.

Molecular dynamics simulation (SD) was used to assess the complex stability of the vaccine protein. In previous studies, macromolecule stability was associated and correlated with the fluctuations of atoms<sup>82,83</sup>. Therefore MD was performed to evaluate the essential dynamics and complex stability of the vaccine based on the protein normal modes in the iMODS server. The analysis showed that no atoms had a significant distortion in the vaccine protein structure indicating less chance of deformability with proper stiffness motion.

It is noteworthy that bioinformatics and immunologic analysis tools provided that the chimeric vaccine should comprises linear and discontinuous B-cell epitopes in addition to MHC-I and MHC-II epitopes<sup>84</sup>. Our predicted vaccine was shown comprising all these epitopes which strongly facilitate the interaction against the humoral and adaptive immunity of the host<sup>84</sup>.

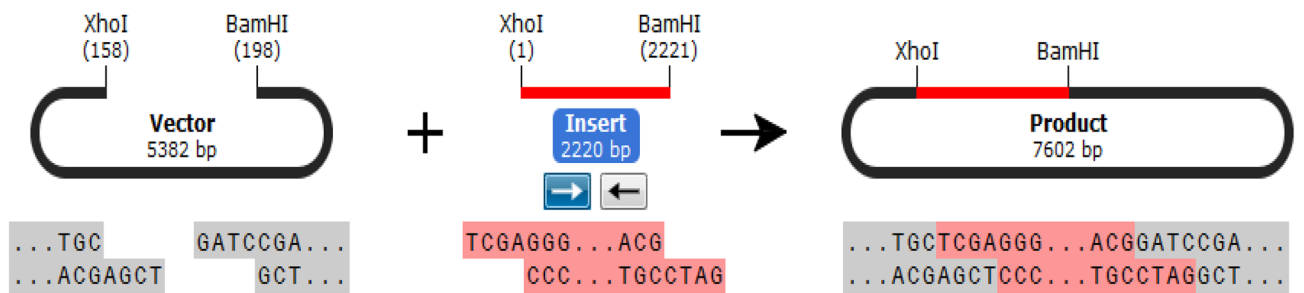
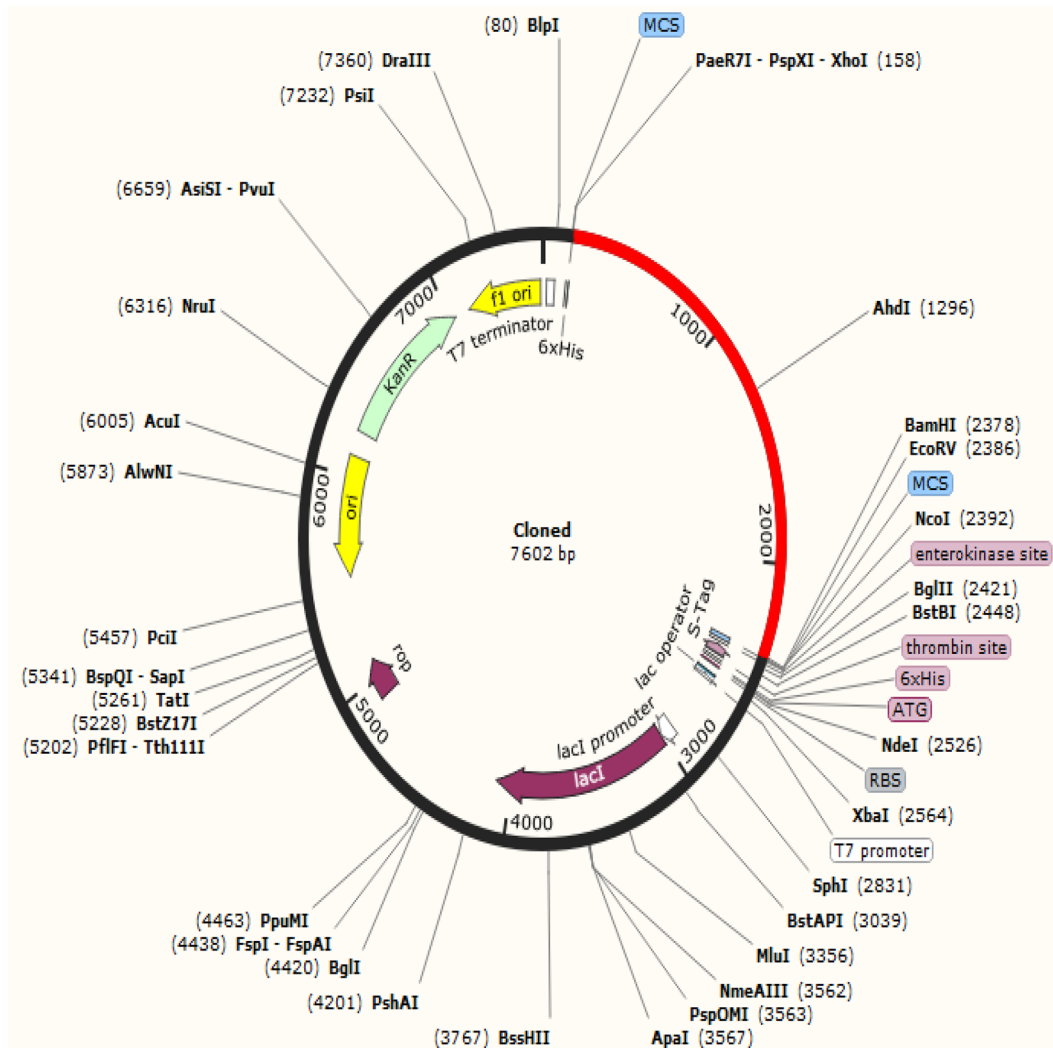
The geometry and topology features of protein structures, such as interior cavities, pockets in the structure surface and the cross channels prior to the docking process are essential to study the function of proteins. The vaccine construct showed surface binding pocket suitable for docking with chicken TLR7. Based on the molecular docking, the constructed vaccine and the TLR7 demonstrated a good binding affinity. The vaccine strongly bound to the chicken TLR7 revealed by the negative values of the docking process<sup>62</sup>. Among the ten chicken TLRs, TLR7 has a propensity to recognize the viral constituents located on the extracellular surfaces<sup>63</sup>, thus has the advantages to be elected for docking against ALV predicted vaccine.

Molecular cloning is an important step to produce recombinant vaccines. Prior to cloning into the pET-30a [+] vector, reverse transcription and adaptation of the vaccine protein to DNA by the JCAT were performed on *E. coli* strain K12. The DNA sequence showed a Cal-Value of 1.00 and a GC ratio of 59.57%, demonstrating a high expression in bacteria. Cloning of the vaccine construct gene in the vector was typically carried out in multiple cloning sites. This result provided prolific cloning of the vaccine protein.

## Conclusion

This study demonstrated the urgency need for effective vaccine strategy against ALV due to the lack of treatment or approved antiviral drugs. This study inclusively exploited the computational and immunoinformatics approaches to design and evaluate a multi-epitope vaccine candidate against ALV. Constructing abjunctive vaccine with antigenic characteristics, devoid of allergenicity and toxicity is a crucial footstep to combat ALV. This study provided a potential vaccine epitopes with immunogenic adjuvant and suitable linkers. The vaccine was stable and provokes strong immune response interactions. Moreover the vaccine showed favorable interaction with the chicken immune receptor as confirmed by molecular docking analysis. However, validation of this vaccine via experimental studies is essential to guarantee the immunogenicity and protective efficacy of the vaccine.





**Figure 10.** The vaccine DNA sequence was cloned in the pET30a (+) vector. The vector was shown in black colour, while the red colour represents the gene coding for the vaccine protein.

### Data availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Received: 16 October 2023; Accepted: 27 January 2024

Published online: 04 February 2024

### References

1. Fadly, A. M. Isolation and identification of avian leukosis viruses: A review. *Avian Pathol.: J. W.V.P.A* **29**(6), 529–535. <https://doi.org/10.1080/03079450020016760> (2000).

2. Fandiño, S., Gomez-Lucia, E., Benítez, L. & Doménech, A. Avian leukosis: Will we be able to get rid of it?. *Anim.: Open Access J. MDPI* **13**(14), 2358. <https://doi.org/10.3390/ani13142358> (2023).
3. Zheng, L. P. *et al.* Current epidemiology and co-infections of avian immunosuppressive and neoplastic diseases in chicken flocks in central China. *Viruses* **14**(12), 2599. <https://doi.org/10.3390/v14122599> (2022).
4. Borodin, A. M. *et al.* Eradication of avian leukosis virus subgroups J and K in broiler cross chickens by selection against infected birds using multilocus PCR. *PLoS One* **17**(6), e0269525. <https://doi.org/10.1371/journal.pone.0269525> (2022).
5. Zavala, G. & Cheng, S. Detection and characterization of avian leukosis virus in Marek's disease vaccines. *Avian Dis.* **50**(2), 209–215. <https://doi.org/10.1637/7444-092405R.1> (2006).
6. Gao, Y. L. *et al.* Avian leukosis virus subgroup J in layer chickens, China. *Emerg. Infect. Dis.* **16**(10), 1637–1638. <https://doi.org/10.3201/eid1610.100780> (2010).
7. Khan, A. *et al.* Lymphoid leukosis in fayoumi birds reared in countryside. *Intl. J. Agric. Biol.* **22**, 620–626. <https://doi.org/10.17957/IJAB/15.1107> (2019).
8. Li, H. *et al.* Aberrant expression of liver microRNA in chickens infected with subgroup J avian leukosis virus. *Virus Res.* **169**(1), 268–271. <https://doi.org/10.1016/j.virusres.2012.07.003> (2012).
9. Payne, L. N. & Nair, V. The long view: 40 years of avian leukosis research. *Avian Pathol: J. W.V.P.A.* **41**(1), 11–19. <https://doi.org/10.1080/03079457.2011.646237> (2012).
10. Abdel-Latif, M. M. & Khalafalla, A. I. Detection by PCR of multiple subgroups of avian leukosis virus (ALV) in broilers in the Sudan. *J. Anim. Vet. Adv.* **4**(3), 407–413 (2005).
11. Li, Q. *et al.* Recombinant subgroup B avian leukosis virus combined with the subgroup J env gene significantly increases its pathogenicity. *Vet. Microbiol.* **250**, 108862. <https://doi.org/10.1016/j.vetmic.2020.108862> (2020).
12. Cui, N., Su, S., Chen, Z., Zhao, X. & Cui, Z. Genomic sequence analysis and biological characteristics of a rescued clone of avian leukosis virus strain JS11C1, isolated from indigenous chickens. *J. Gen. Virol.* **95**(Pt 11), 2512–2522. <https://doi.org/10.1099/vir.0.067264-0> (2014).
13. Parvizpour, S., Pourseif, M. M., Razmara, J., Rafi, M. A. & Omid, Y. Epitope-based vaccine design: A comprehensive overview of bioinformatics approaches. *Drug Discov. Today* **25**(6), 1034–1042. <https://doi.org/10.1016/j.drudis.2020.03.006> (2020).
14. Li, H. *et al.* Diversity of avian leukosis virus subgroup J in local chickens, Jiangxi, China. *Sci. Rep.* **11**(1), 4797. <https://doi.org/10.1038/s41598-021-84189-7> (2021).
15. Kheimar, A. *et al.* A genetically engineered commercial chicken line is resistant to highly pathogenic avian leukosis virus subgroup. *J. Microorgan.* **9**(5), 1066. <https://doi.org/10.3390/microorganisms9051066> (2021).
16. Payne, L. N. *et al.* A novel subgroup of exogenous avian leukosis virus in chickens. *J. Gen. Virol.* **72**(Pt 4), 801–807. <https://doi.org/10.1099/0022-1317-72-4-801> (1991).
17. Wang, X., Zhao, P. & Cui, Z. Z. Identification of a new subgroup of avian leukosis virus isolated from Chinese indigenous chicken breeds]. *Bing Du XueBao* **28**(6), 609–614 (2012) (Chinese).
18. Patronov, A. & Doytchinova, I. T-cell epitope vaccine design by immunoinformatics. *Open Biol.* **3**(1), 120139. <https://doi.org/10.1098/rsob.120139> (2013).
19. Vij, S., Thakur, R. & Rishi, P. Reverse engineering approach: A step towards a new era of vaccinology with special reference to Salmonella. *Exp. Rev. Vacc.* **21**(12), 1763–1785. <https://doi.org/10.1080/14760584.2022.2148661> (2022).
20. Zhang, J. *et al.* The immunogenicity and protective immunity of multi-epitopes DNA prime-protein boost vaccines encoding Amastin-Kmp-11, Kmp11-Gp63 and Amastin-Gp63 against visceral leishmaniasis. *PLoS One* **15**(3), e0230381. <https://doi.org/10.1371/journal.pone.0230381> (2020).
21. Saldanha, L., Langel, Ü. & Vale, N. In Silico studies to support vaccine development. *Pharmaceutics* **15**(2), 654. <https://doi.org/10.3390/pharmaceutics15020654> (2023).
22. Jorge, S. & Dellagostin, O. A. The development of veterinary vaccines: A review of traditional methods and modern biotechnology approaches. *Biotechnol. Res. Innov.* **1**, 6–13. <https://doi.org/10.1016/j.biori.2017.10.001> (2017).
23. Silva-Arrieta, S., Goulder, P. J. R. & Brander, C. In silico veritas? Potential limitations for SARS-CoV-2 vaccine development based on T-cell epitope prediction. *PLoS Pathog.* **16**(6), e1008607. <https://doi.org/10.1371/journal.ppat.1008607> (2020).
24. Shawan, M. M. A. K. *et al.* Advances in computational and bioinformatics tools and databases for designing and developing a multi-epitope-based peptide vaccine. *Int. J. Pept. Res. Ther.* **29**(4), 60. <https://doi.org/10.1007/s10989-023-10535-0> (2023).
25. Dey, J. *et al.* Designing a novel multi-epitope vaccine to evoke a robust immune response against pathogenic multidrug-resistant *Enterococcus faecium* bacterium. *Gut Pathog.* **14**(1), 21. <https://doi.org/10.1186/s13099-022-00495-z> (2022).
26. NCBI Resource Coordinators. Database resources of the National Center for Biotechnology Information. *Nucl. Acids Res.* **44**(D1), D7–D19. <https://doi.org/10.1093/nar/gkv1290> (2016).
27. Gasteiger, E., Hoogland, C., Gattiker, A., Wilkins, M. R., Appel, R.D. & Bairoch, A. *Protein Identification and Analysis Tools on the ExPASy Server. The Proteomics Protocols Handbook*, pp. 571–607 (2005). <https://doi.org/10.1385/1-59259-890-0:571>
28. Doytchinova, I. A. & Flower, D. R. Vaxijen: A server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinform.* **8**, 4. <https://doi.org/10.1186/1471-2105-8-4> (2007).
29. Zaharieva, N., Dimitrov, I., Flower, D. & Doytchinova, I. Immunogenicity prediction by VaxiJen: A 10 year overview. *J. Proteom. Bioinform.* <https://doi.org/10.4172/jpb.1000454> (2017).
30. Dimitrov, I., Bangov, I., Flower, D. R. & Doytchinova, I. AllerTOP vol 2—a server for prediction of allergens. *J. Mol. Model.* **20**(6), 2278. <https://doi.org/10.1007/s00894-014-2278-5> (2014).
31. Gupta, S., Open Source Drug Discovery Consortium *et al.* Approach for predicting toxicity of peptides and proteins. *PLoS One* **8**(9), e73957. <https://doi.org/10.1371/journal.pone.0073957> (2013).
32. Mao, Y., Su, Q., Li, J., Jiang, T. & Wang, Y. Avian leukosis virus contamination in live vaccines: A retrospective investigation in China. *Vet. Microbiol.* **246**, 108712. <https://doi.org/10.1016/j.vetmic.2020.108712> (2020).
33. Käll, L., Krogh, A. & Sonnhammer, E. L. Advantages of combined transmembrane topology and signal peptide prediction—the Phobius web server. *Nucl. Acids Res.* **35**(Web Server issue), W429–W432. <https://doi.org/10.1093/nar/gkm256> (2007).
34. Hall, T. A. BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. in *Nucleic Acids Symposium Series*, Jan 1 (Vol. 41, No. 41, pp. 95–98). Information Retrieval Ltd., c1979–c2000 (1999)
35. Ansari, H. R. & Raghava, G. P. Identification of conformational B-cell epitopes in an antigen from its primary sequence. *Immun. Res.* **6**, 6. <https://doi.org/10.1186/1745-7580-6-6> (2010).
36. Saha, S. & Raghava, G. P. Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. *Proteins* **65**(1), 40–48. <https://doi.org/10.1002/prot.21078> (2006).
37. Vita, R. *et al.* The immune epitope database (IEDB): 2018 update. *Nucl. Acids Res.* **47**(D1), D339–D343. <https://doi.org/10.1093/nar/gky1006> (2019).
38. Dhanda, S. K. *et al.* IEDB-AR: Immune epitope database-analysis resource in 2019. *Nucl. Acids Res.* **47**(W1), W502–W506. <https://doi.org/10.1093/nar/gkz452> (2019).
39. Mugunthan, S. P. & Harish, M. C. Multi-epitope-based vaccine designed by targeting cytoadherence proteins of mycoplasma gal-lisepticum. *ACS Omega* **6**(21), 13742–13755. <https://doi.org/10.1021/acsomega.1c01032> (2021).
40. Andongma, B. T. *et al.* design of a promiscuous chimeric multi-epitope vaccine against *Mycobacterium tuberculosis*. *Comput. Struct. Biotechnol. J.* **21**, 991–1004. <https://doi.org/10.1016/j.csbj.2023.01.019> (2023).



41. Nielsen, M. & Lund, O. NN-align An artificial neural network-based alignment algorithm for MHC class II peptide binding prediction. *BMC Bioinform.* **10**, 296. <https://doi.org/10.1186/1471-2105-10-296> (2009).
42. Omoniyi, A. A. *et al.* design and analyses of a multi-epitope vaccine against Crimean-Congo hemorrhagic fever virus through reverse vaccinology and immunoinformatics approaches. *Sci. Rep.* **12**(1), 8736. <https://doi.org/10.1038/s41598-022-12651-1> (2022).
43. Pandey, R. K., Ojha, R., Aathmanathan, V. S., Krishnan, M. & Prajapati, V. K. Immunoinformatics approaches to design a novel multi-epitope subunit vaccine against HIV infection. *Vaccine* **36**(17), 2262–2272. <https://doi.org/10.1016/j.vaccine.2018.03.042> (2018).
44. Almofti, Y. A., Abd-Elrahman, K. A. & Eltilib, E. E. M. Vaccinomic approach for novel multi epitopes vaccine against severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2). *BMC Immunol.* **22**(1), 22. <https://doi.org/10.1186/s12865-021-00412-0> (2021).
45. Mahmoud, N. A., Elshafei, A. M. & Almofti, Y. A. A novel strategy for developing vaccine candidate against Jaagsiekte sheep retrovirus from the envelope and gag proteins: An in-silico approach. *BMC Vet. Res.* **18**(1), 343. <https://doi.org/10.1186/s12917-022-03431-0> (2022).
46. Wang, S., Li, W., Liu, S. & Xu, J. RaptorX-Property: A web server for protein structure property prediction. *Nucl. Acids Res.* **44**(W1), W430–W435. <https://doi.org/10.1093/nar/gkw306> (2016).
47. Ko, J., Park, H., Heo, L. & Seok, C. GalaxyWEB server for protein structure prediction and refinement. *Nucl. Acids Res.* **40**(Web Server issue), W294–W297. <https://doi.org/10.1093/nar/gks493> (2012).
48. Seok, C. *et al.* Accurate protein structure prediction: What comes next?. *Bio Des.* **9**, 47–50. <https://doi.org/10.34184/kssb.2021.9.3.47> (2021).
49. Srivastava, S. *et al.* Structural Basis for designing multi-epitope vaccines against COVID-19 infection: Vaccine design and validation. *JMIR Bioinform. Biotechnol.* **1**(1), e19371. <https://doi.org/10.2196/19371> (2020).
50. Lovell, S. C. *et al.* Structure validation by Calpha geometry: phi, psi and Cbeta deviation. *Proteins* **50**(3), 437–450. <https://doi.org/10.1002/prot.10286> (2003).
51. Wiederstein, M. & Sippl, M. ProSA-web: Interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucl. Acids Res.* **35**, W407–W410. <https://doi.org/10.1093/nar/gkm290> (2007).
52. Shawan, M. M. A. K. *et al.* Insilico modeling and immunoinformatics probing disclose the epitope based peptide vaccine against Zika virus envelope glycoprotein. *Indian J. Pharm. Biol. Res.* **2**, 44–57. <https://doi.org/10.30750/ijpbr.2.4.10> (2014).
53. Hebditch, M., Carballo-Amador, M. A., Charonis, S., Curtis, R. & Warwicker, J. Protein-Sol: A web tool for predicting protein solubility from sequence. *Bioinformatics (Oxford, England)* **33**(19), 3098–3100. <https://doi.org/10.1093/bioinformatics/btx345> (2017).
54. Magnan, C. N., Randall, A. & Baldi, P. SOLpro: Accurate sequence-based prediction of protein solubility. *Bioinformatics (Oxford, England)* **25**(17), 2200–2207. <https://doi.org/10.1093/bioinformatics/btp386> (2009).
55. Craig, D. B. & Dombkowski, A. A. Disulfide by design 2.0: A web-based tool for disulfide engineering in proteins. *BMC Bioinform.* **14**, 346. <https://doi.org/10.1186/1471-2105-14-346> (2013).
56. Rapin, N., Lund, O., Bernaschi, M. & Castiglione, F. Computational immunology meets bioinformatics: The use of prediction tools for molecular binding in the simulation of the immune system. *PLoS One* **5**(4), e9862. <https://doi.org/10.1371/journal.pone.0009862> (2010).
57. López-Blanco, J. R., Aliaga, J. I., Quintana-Ortí, E. S. & Chacón, P. iMODS: Internal coordinates normal mode analysis server. *Nucl. Acids Res.* **42**(Web Server issue), W271–W276. <https://doi.org/10.1093/nar/gku339> (2014).
58. López-Blanco, J. R., Garzón, J. I. & Chacón, P. iMod: multipurpose normal mode analysis in internal coordinates. *Bioinformatics (Oxford, England)* **27**(20), 2843–2850. <https://doi.org/10.1093/bioinformatics/btr497> (2011).
59. Ponomarenko, J. *et al.* ElliPro: A new structure-based tool for the prediction of antibody epitopes. *BMC Bioinform.* **9**, 514. <https://doi.org/10.1186/1471-2105-9-514> (2008).
60. Tian, W., Chen, C., Lei, X., Zhao, J. & Liang, J. CASTp 3.0: Computed atlas of surface topography of proteins. *Nucl. Acids Res.* **46**(W1), W363–W367. <https://doi.org/10.1093/nar/gky473> (2018).
61. Halder, S. K. *et al.* In silico identification and analysis of potentially bioactive antiviral phytochemicals against SARS-CoV-2: A molecular docking and dynamics simulation approach. *BioMed Res. Int.* **2023**, 5469258. <https://doi.org/10.1155/2023/5469258> (2023).
62. Honorato, R. V. *et al.* Structural biology in the clouds: The WeNMR-EOSC ecosystem. *Front. Mol. Biosci.* **8**, 729513. <https://doi.org/10.3389/fmolb.2021.729513> (2021).
63. Vinkler, M., Bainová, H. & Bryja, J. Protein evolution of Toll-like receptors 4, 5 and 7 within Galloanserae birds. *Genet. Select. Evolut.: GSE* **46**(1), 72. <https://doi.org/10.1186/s12711-014-0072-6> (2014).
64. Vangone, A. & Bonvin, A. M. Contacts-based prediction of binding affinity in protein-protein complexes. *eLife* **4**, e07454. <https://doi.org/10.7554/eLife.07454> (2015).
65. Xue, L. C., Rodrigues, J. P., Kastiritis, P. L., Bonvin, A. M. & Vangone, A. PRODIGY: A web server for predicting the binding affinity of protein-protein complexes. *Bioinformatics (Oxford, England)* **32**(23), 3676–3678. <https://doi.org/10.1093/bioinformatics/btw514> (2016).
66. Laskowski, R. A., Jabłońska, J., Pravda, L., Vařeková, R. S. & Thornton, J. M. PDBsum: Structural summaries of PDB entries. *Prot. Sci.: Publ. Prot. Soc.* **27**(1), 129–134. <https://doi.org/10.1002/pro.3289> (2018).
67. Grote, A. *et al.* JCat: A novel tool to adapt codon usage of a target gene to its potential expression host. *Nucl. Acids Res.* **33**(Web Server issue), W526–W531. <https://doi.org/10.1093/nar/gki376> (2005).
68. Ravikumar, R., Chan, J. & Prabaharan, M. Vaccines against major poultry viral diseases: Strategies to Improve the breadth and protective efficacy. *Viruses* **14**(6), 1195. <https://doi.org/10.3390/v14061195> (2022).
69. Feng, M. & Zhang, X. Immunity to Avian Leukosis Virus: Where are we now and what should we do?. *Front. Immunol.* **7**, 624. <https://doi.org/10.3389/fimmu.2016.00624> (2016).
70. Xu, Q., Ma, X., Wang, F., Li, H. & Zhao, X. Evaluation of a multi-epitope subunit vaccine against avian leukosis virus subgroup J in chickens. *Virus Res.* **210**, 62–68. <https://doi.org/10.1016/j.virusres.2015.06.024> (2015).
71. Xu, Q. *et al.* Design and construction of a chimeric multi-epitope gene as an epitope-vaccine strategy against ALV-J. *Prot. Expr. Purif.* **106**, 18–24. <https://doi.org/10.1016/j.pep.2014.10.007> (2015).
72. Wang, X. *et al.* A novel multi-variant epitope ensemble vaccine against avian leukosis virus subgroup J. *J. Vacc.* **35**(48 Pt B), 6685–6690. <https://doi.org/10.1016/j.vaccine.2017.10.019> (2017).
73. Wang, S. *et al.* Oral immunization of chickens with recombinant *Lactobacillus plantarum* vaccine against early ALV-J infection. *Front. Immunol.* **10**, 2299. <https://doi.org/10.3389/fimmu.2019.02299> (2019).
74. Sun, Y. *et al.* Antiviral activity against avian leucosis virus subgroup J of degraded polysaccharides from *Ulva pertusa*. *BioMed Res. Int.* **9415965**, 2018. <https://doi.org/10.1155/2018/9415965> (2018).
75. Barman, A., Deb, B. & Chakraborty, S. Prediction of potential epitopes for peptide vaccine formulation against teschovirus a using immunoinformatics. *Int. J. Pept. Res. Ther.* **26**(2), 1137–1146. <https://doi.org/10.1007/s10989-019-09916-1> (2020).
76. Kar, T. *et al.* A candidate multi-epitope vaccine against SARS-CoV-2. *Sci. Rep.* **10**(1), 10895. <https://doi.org/10.1038/s41598-020-67749-1> (2020).
77. Zepp, F. Principles of vaccine design-Lessons from nature. *Vaccine* **28**(Suppl 3), C14–C24. <https://doi.org/10.1016/j.vaccine.2010.07.020> (2010).

78. Ojha, R., Pareek, A., Pandey, R. K., Prusty, D. & Prajapati, V. K. Strategic development of a next-generation multi-epitope vaccine to prevent Nipah virus zoonotic infection. *ACS Omega* **4**(8), 13069–13079. <https://doi.org/10.1021/acsomega.9b00944> (2019).
79. Nain, Z. *et al.* Proteome-wide screening for designing a multi-epitope vaccine against emerging pathogen *Elizabethkingia anophelis* using immunoinformatic approaches. *J. Biomol. Struct. Dyn.* **38**(16), 4850–4867. <https://doi.org/10.1080/07391102.2019.1692072> (2020).
80. Fan, J. *et al.* Engineering his-tagged senecavirus a for one-step purification of viral antigens. *Vaccines* **10**(2), 170. <https://doi.org/10.3390/vaccines10020170> (2022).
81. Landman, W. J. *et al.* Effect of an in ovo infection with a Dutch avian leukosis virus subgroup J isolate on the growth and immunological performance of SPF broiler chickens. *Avian Pathol: J. W.V.P.A* **31**(1), 59–72. <https://doi.org/10.1080/03079450120106633> (2002).
82. Cheatham, T. I., Miller, J. L., Fox, T., Darden, T. A. & Kollman, P. A. Molecular dynamics simulations on solvated biomolecular systems: the particle mesh Ewald method leads to stable trajectories of DNA, RNA, and proteins. *J. Am. Chem. Soc.* **117**(14), 4193–4194. <https://doi.org/10.1021/ja00119a045> (1995).
83. Tehei, M. & Zaccai, G. Adaptation to high temperatures through macromolecular dynamics by neutron scattering. *FEBS J.* **274**(16), 4034–4043. <https://doi.org/10.1111/j.1742-4658.2007.05953.x> (2007).
84. Ali, M. *et al.* Exploring dengue genome to construct a multi-epitope based subunit vaccine by utilizing immunoinformatics approach to battle against dengue infection. *Sci. Rep.* **7**(1), 9232. <https://doi.org/10.1038/s41598-017-09199-w> (2017).

## Acknowledgements

Special acknowledge to Lamees Tarig Mohamed for writing assistance

## Author contributions

S.O.E.: conceived the study, designed the study and wrote and drafted the manuscript. N.A.M.: conceived the study, designed the study, did data analysis, and wrote and drafted the manuscript. Y.A.A.: Conceived the study and data analysis, coordinated the study, and drafted the manuscript. All authors gave final approval for publication.

## Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors Ethics declarations.

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to Y.A.A.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024