processes

MDPI

Article

In Silico Epitope-Based Peptide Vaccine Design Against Influenza B Virus: An Immunoinformatics Approach

Hao Wu [1,2], Chenyan Zhao [1], Ziqi Cheng [1], Weijin Huang [1,3,*] and Yongxin Yu [1,*]

[1] Institute for Biological Product Control, National Institutes for Food and Drug Control (NIFDC), Beijing 102629, China; wh1031420724@163.com (H.W.); zhaochenyan@nifdc.org.cn (C.Z.); chengzq21@mails.jlu.edu.cn (Z.C.)
[2] The Second Laboratory of Viral Vaccine Research, Wuhan Institute of Biological Products Co., Ltd., Wuhan 430207, China
[3] Chinese Academy of Medical Science & Peking Union Medical College, Beijing 100730, China
* Correspondence: huangweijin@nifdc.org.cn (W.H.); yuyongxin@nifdc.org.cn (Y.Y.)

Abstract: Background/Objectives: Influenza viruses are highly transmissible and mutable, posing a significant burden on public health. This study aimed to design a recombinant multi-epitope vaccine with broad protective potential. Methods: Immunoinformatic approaches were employed to predict epitopes from over 30,000 protein sequences retrieved from protein databases. Epitopes were filtered using four key indicators: antigenicity, allergenicity, toxicity, and conservancy. Population coverage analysis was conducted to estimate the proportion of the global population that could potentially benefit from the vaccine. Secondary and tertiary structures of the recombinant vaccine were predicted using the PSIPRED server and AlphaFold2. The vaccine efficacy was validated through an immune simulation, molecular docking, and molecular dynamics simulation. Results: A recombinant multi-epitope vaccine demonstrating strong antigenicity, no allergenicity or toxicity, and high conservation across different subtypes was successfully constructed. Population coverage analysis indicated that the vaccine could elicit an immune response in 90.14% of the global population. Both the secondary and tertiary structures of the vaccine were accurately predicted. Molecular dynamics simulations further validated the structural stability and interactions of the vaccine components with TRL4. Molecular docking confirmed the robust binding affinity of T-cell epitopes to MHC molecules. Simulated immunity studies showed that the vaccine induced the proliferation of memory B cells and T cells, enabling rapid antibody production during viral challenges. Conclusions: This study provides a promising basis for the development of a broadly protective influenza vaccine, leveraging cutting-edge immunoinformatics and molecular dynamics simulations to address the global challenge posed by influenza virus variability.

Keywords: influenza B; immunoinformatic; structure prediction; recombinant vaccine; immune simulation

check for updates

Academic Editor: Prihardi Kahar

Received: 26 December 2024
Revised: 11 February 2025
Accepted: 23 February 2025
Published: 27 February 2025

Citation: Wu, H.; Zhao, C.; Cheng, Z.; Huang, W.; Yu, Y. In Silico Epitope-Based Peptide Vaccine Design Against Influenza B Virus: An Immunoinformatics Approach. *Processes* **2025**, *13*, 681. https://doi.org/10.3390/pr13030681

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

The spread of influenza viruses in humans and the heavy burden it imposes on public health has been widely described [1–3]. Based on relevant epidemiological studies, about a quarter of the 1.8 million influenza cases from 2000 to 2018 were caused by influenza B viruses (IBV) [3]. Especially in pediatric patients, IBV may cause more severe illness than influenza A viruses [4–6]. An analysis of influenza surveillance data in the United States for the period 2004–2019 showed that a significant proportion (7–51%) of pediatric

deaths with confirmed influenza were associated with IBV [7]. Although IBV are not a major factor in influenza pandemics, the threat they pose to human health should not be underestimated [8]. IBV are classified as a single influenza type, but two antigenically and genetically distinct lineages are transmitted together, represented by B/Victoria/2/1987 (Victoria lineage) (BV) and B/Yamagata/16/1988 (Yamagata lineage) (BY) [9,10]. Both lineages have shown unusually high levels of epidemiological activity since 2015, with hemagglutinin (HA)-deficient variants appearing simultaneously in multiple Victorian lineages [10]. At the early stages of the COVID-19 pandemic (starting in March 2020), and even after the flu virus began circulating again in late 2021, the B/Yamagata lineage was detected in only a few countries, raising concerns about its potential extinction. However, it is premature to declare the lineage extinct, as transmission may currently be at low levels (below surveillance detection thresholds) or in regions not fully monitored, leaving room for its potential resurgence [11].

The influenza virus genome consists of eight segments of single-stranded RNA encoding at least eleven proteins [12], including the two major surface glycoproteins HA and neuraminidase (NA) [13], matrix protein 1 and 2 (M1 and M2), two non-structural proteins (NS1 and NS2), nucleoprotein (NP), and four polymerase-associated proteins (PB1, PB2, PB1-F2, and PA). HA is a homotrimer derived from its precursor protein HA0 by cleavage, folding, and modification [14,15]. Most immune responses induced by vaccines or natural infections are directed against the HA receptor binding domain (R domain) [16]. Since the R domain of HA is highly diverse among strains, vaccines tend to elicit protection only against the same subtype of influenza virus [17]. NA protein is a homotetrameric type II membrane protein shaped like a mushroom and contains approximately 470 residues per monomer [18]. Natural infection triggers an immune response not only against HA, but also against NA, which is equally important. The fact that NA has less variation than HA makes its protein sequences highly conserved across strains [19]. M1 is one of the most abundant and conserved proteins in influenza viruses [20], and it is also a potential target for a universal vaccine.

One of the most urgent tasks for the next decade, as announced by The World Health Organization (WHO) in "WHO Preferred Product Characteristics for Next-Generation Influenza Vaccines" is the manufacture of vaccines that provide long-term protection against severe and complicated forms of influenza A and B [21]. Many organizations or institutions are working to develop such a universal vaccine, but the focus of the research is on the influenza A virus [22]. It is clear that the antigenic variation of IBV is not as extensive as that of influenza A viruses. However, the need for a universal IBV vaccine is equally urgent. Currently, IBV is not used as a single antigenic component to construct vaccines, but is recommended to be used together with the influenza A virus in the construction of trivalent or quadrivalent vaccines [23]. Historically, seasonal influenza vaccine formulations have contained two IAV strains of the H3N2 and H1N1 subtypes, and one IBV strain from one of the two lineages. Alternately, both lineages can be used together to make a quadrivalent vaccine [24]. Meanwhile, the WHO and local governments give recommendations on the composition of influenza vaccines based on surveillance data from the previous season [23,25]. This raises two main problems: the mismatch between the vaccine strain and the prevalent strain, and the need to continuously update the vaccine components on an annual basis. In recent years, predictions of the major influenza B strains have been very inaccurate. From 2001–2002 to 2010–2011, the prevalence of the major influenza B strains differed from vaccine strains in 5 out of 10 influenza seasons in the United States. During this period, 46% of influenza B samples were of IBV strains not included in the vaccine [26]. Data from Bangladesh show multiple branches of influenza strains for each season between 2009 and 2020 with significant mismatches

to WHO-recommended vaccine strains [27]. Such mismatches have occurred in several countries and have been proven to directly reduce the protective effect of vaccines [28,29].

Based on these phenomena, the ultimate goal of influenza vaccine research is to design immunogens that are against not only current strains, but also future strains that mutate due to antigenic drift and/or antigenic shift [30]. Updated vaccine platforms are needed that have the potential to induce humoral and cellular immune responses and thus protect against a broad range of influenza viruses [31]. Protective immunity against the influenza virus relies primarily on adaptive immune responses. Neutralizing antibodies, especially those targeting hemagglutinin (HA) and neuraminidase (NA), block viral entry and promote clearance. $CD4^+$ T cells support antibody production, while $CD8^+$ cytotoxic T cells directly eliminate infected cells. These adaptive mechanisms are crucial for effective viral clearance and long-term protection. With the advent and development of emerging disciplines such as bioinformatics, structural biology, and genomics, new vaccine design strategies such as reverse vaccinology approaches, serum vaccinology approaches, and analytical vaccinology approaches provide new and feasible paths for the development of universal vaccines. By analyzing viral genomic sequences and the structures of encoded antigens, immunoinformatics can rapidly predict the impact of mutations on antigenicity and immune cross-protection. Additionally, through machine learning, molecular dynamics (MD) simulations, and epitope prediction algorithms, it identifies the most broadly protective and effective immunogenic epitopes, guiding the development of multivalent vaccines. Different approaches have been used, including the rational design of the HA stem antigen, a recombinant adenovirus vaccine carrying the consensus sequence of all IBV hemagglutinins, mosaic hemagglutinin-based whole inactivated virus vaccines, etc. All these vaccine candidates have been proven to protect against homologous or heterologous challenges to some extent.

In this study, we utilized immunoinformatics and structural biology tools to design an epitope-based peptide vaccine targeting HA, NA, and M1 proteins. Our approach aims to address the challenges of strain mismatches and provide broad protection against IBV. To achieve this goal, we collected the protein sequences from 1993 to 2022 and calculated their common sequences separately. Subsequently, B-cell epitopes and T-cell epitopes were predicted by immunoinformatics on a common sequence to construct tandem epitope antigens. AlphaFold2 was used to model the recombinant antigen, and then the protective effect of the vaccine was validated by computer simulation methods.

2. Materials and Methods

Recombinant vaccine with tandem epitopes was constructed by immunoinformatics techniques and its protective effect was verified by a computer simulation with a tertiary structure prediction by AlphaFold2 (ColabFold v1.5.5: AlphaFold2 using MMseqs2). All experiments were conducted on a personal computer and used various online and offline tools. The flowchart in Figure 1 depicts the workflow of this section.

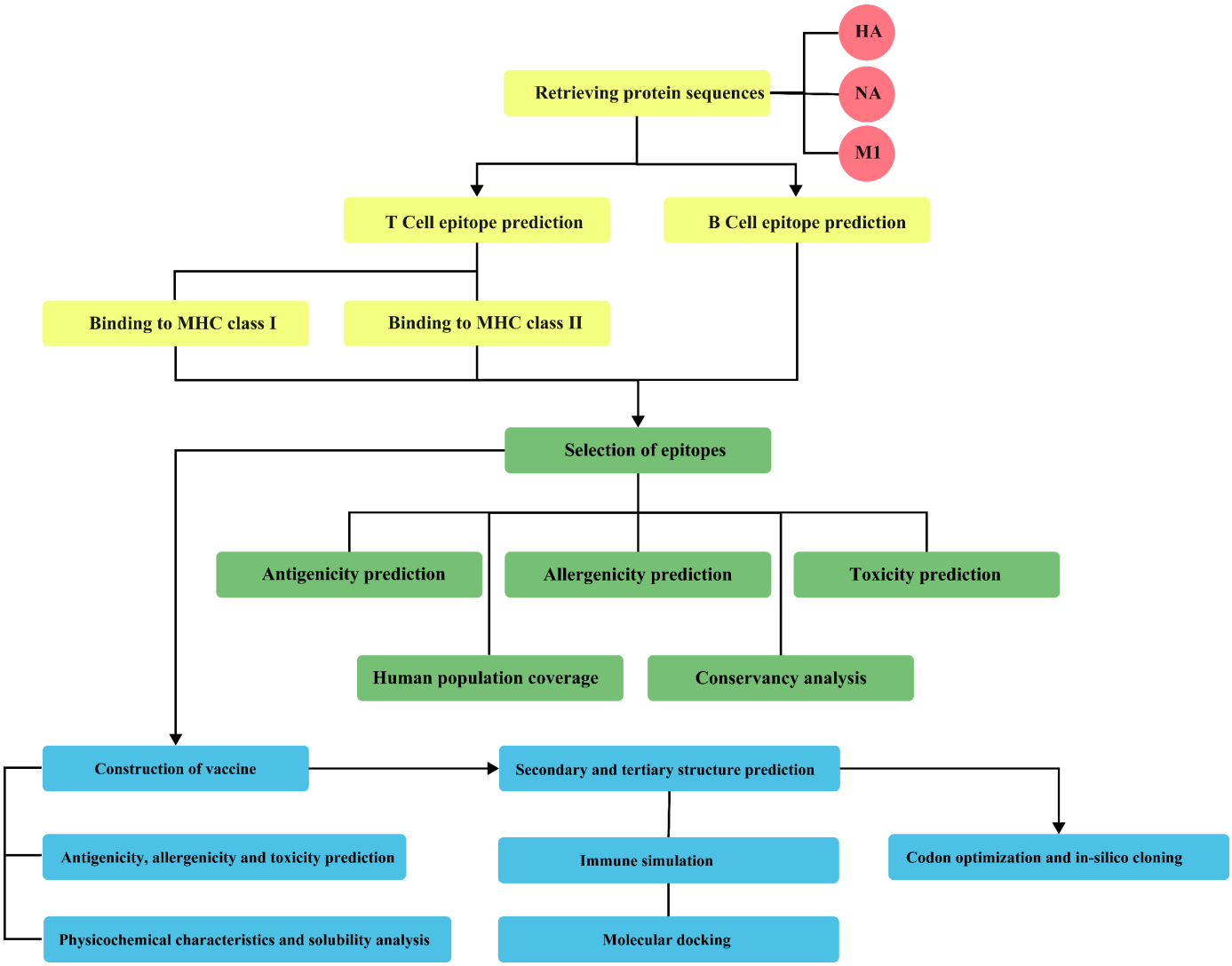

Figure 1. Flowchart of multi-epitope vaccine construction.

2.1. Retrieving Protein Sequences of IBV and Multiple Alignments

The protein database in the National Center for Biotechnology Information (NCBI) was used to select and obtain information on the three proteins, HA, NA and M1. The species was restricted to IBV, and the precursor proteins and fragments were shielded by a sequence length restriction. The complete sequences of all proteins were downloaded in FASTA format without restricting the lineage to BV or BY.

Multiple sequence alignment was performed using the Super5 algorithm of the Muscle v5 program on a local system [32]. Super5 algorithm was designed to scale parallel perturbed probcons by introducing divide-and-conquer heuristics for the alignment of large-scale sequences. The aligned sequences were imported into Jalview v2.11.4.1 to extract the common sequence [33].

2.2. T-Cell Epitope Prediction

T-cell epitope predictions were performed in Immune Epitope Database (IEDB) and Services in DTU Health Tech using common sequence. These tools predict IC50 values for peptides binding to specific MHC molecules.

2.2.1. Peptide Binding to MHC Class I Molecules

NetMHCPan v4.1 method was used for MHC class I binding prediction [34]. The experiment restricted the MHC source species to human and used the IEDB-recommended allele reference set containing 27 alleles (Table S1). The length of the peptide was 9–10 m. The predicted output was sorted by IC50 nM. Peptides with IC values < 50 nM was considered high affinity and selected as candidate epitopes.

2.2.2. Peptide Binding to MHC Class II Molecules

IEDB recommended method was used for MHC class II binding prediction [35]. Similar to MHC I predictions, MHC II analyses were restricted to human variation using the complete HLA reference set of 27 alleles (Table S1). For each peptide, a percentile rank for each of the three methods (combinatorial library, SMM-align, and Sturniolo) is generated by comparing the peptide's score against the scores of five million random 15 m selected from SWISSPROT database. And the adjusted percentile rank is adjusted for the frequency of peptide lengths. The median percentile rank of the three methods were then used to generate the rank for consensus method. The epitopes ranked in the top five percent were selected as candidates.

2.3. B-Cell Epitope Prediction

Three different methods were used for the prediction of B-cell epitopes. Emini surface accessibility scale was based on surface accessibility scale of the input sequence [36]. Kolaskar and Tongaonkar antigenicity scale was constructed using mobility of protein segments based on the temperature B factors of 31 proteins with known structure [37]. BepiPred v2.0 used a random forest algorithm trained on epitopes and non-epitope amino acids determined from crystal structures [38]. Duplicates of the epitopes predicted by the three methods were removed, and the remaining were used as candidate B-cell epitopes.

2.4. Antigenicity, Allergenicity, and Toxicity Prediction of Epitopes

Each predicted epitope was examined for antigenicity, allergenicity, and toxicity.

VaxiJen v2.0 server was used for antigenicity assessment of epitopes [39,40]. The target species was selected as the virus, and the threshold of the model was set to 0.4, meaning that epitopes greater than this value were considered antigenic.

Aller TOP v2.0 server and Allergen FP v1.0 server were used for allergenicity assessment of epitopes [41,42]. Epitopes determined to be non-allergenic in both assessments were considered to pass the allergenicity test. In humoral immunity, immunity against HA should predominate in natural infections. Therefore, a search for suitable candidates was renewed among the B-cell epitopes of HA eliminated in the previous screening step. In terms of screening requirements, the restrictions were relaxed only in terms of allergenicity, and the new criteria allow epitopes with non-allergenic results in only one allergenicity test to become candidate epitopes again. In terms of HA screening requirements, only allergenicity was relaxed, and the new criteria allowed epitopes with no allergen results in only one allergen test to be candidate epitopes.

ToxinPred server was used to evaluate the toxicity of epitopes using the SVM (Swiss-Prot) method with the threshold set to 0 [43,44].

2.5. Conservancy and Human Population Coverage Analysis

Conservation of epitopes was predicted using the Epitope Conservancy Analysis tool [45,46], an extension of the IEDB database. All epitopes conserved in more than 50% of the strains were retained. Population coverage was calculated worldwide using a combined

Class I and II mode. T cell epitopes eligible for the previous step were examined using the population coverage tool [47].

2.6. Construction of Vaccine

Candidate epitopes with strong antigenicity, non-allergenicity, non-toxicity, and broad protective effects were used for vaccine design, as showed in Table S5. Two adjuvants (human β-defensin-3 and LL37) and four protein linkers (EAAAK, GGGS, KK, and AAY) were selected for the composition of vaccine [48,49].

2.7. Antigenicity, Allergenicity, and Toxicity Prediction of Recombinant Vaccine

Antigenicity and allergenicity of recombinant vaccine were examined using the same method as in 2.4. Toxicity was evaluated using the ToxinPred2 server [50]. ToxinPred2 is developed for predicting toxicity of proteins. After entering the protein sequence of the vaccine, hybrid technique was selected to develop the model and the threshold was set to 0.6.

2.8. Physicochemical Characteristics and Solubility of Recombinant Vaccine

Solubility of vaccine peptides was predicted using the protein-sol server [51], while physicochemical characteristics, including molecular weight, theoretical pI, estimated half-life, instability index, aliphatic index, and grand average of hydropathicity (GRAVY) were obtained using the ProtParam server [52,53].

2.9. Tertiary Structure Prediction, Energy Minimization, and Validation

AlphaFold2 was used for tertiary structure modeling of vaccine peptides and T-cell epitopes [54]. Tertiary structures were modeled in AlphaFold2 using the monomer model, while running with all genetic databases used at CASP14. The structure was improved by energy minimization using the Galaxy server and then validated using two severs: ERRAT and ProCheck Ramachandran plots.

2.10. Molecular Docking

Molecular docking was calculated using the ZDOCK program in Discovery Studio Workstation. Blind docking of rigid molecules was used during docking, and the angular step size was set to 15. The vaccine constructs and T-cell epitopes were used as ligands for docking. TRL4 (PDB ID: 3FXI) and six different MHC molecules were chosen for the receptor proteins, including two MHC class I molecules, HLA-A*03:01 (PDB ID: 7MLE) and HLA-B*27:05 (PDB ID: 5IB1), and four MHC class II molecules, DRB1*0101 (PDB ID: 2FSE), DRB3*0101 (PDB ID: 2Q6W), HLA-DPA1*02:01 (PDB ID: 3LQZ), and HLA-DQA1*05:01 (PDB ID: 1JK8). The ligand conformation in the antigen-binding groove was selected, further optimized by RDOCK, and performed a re-scoring to determine the final docking structure.

2.11. Molecular Dynamics Simulation

Molecular dynamics simulations were conducted using GROMACS 2022.3 software. The simulations were performed under constant temperature (310.15 K) and atmospheric pressure (1 Bar). The Amber99SB-ILDN force field was employed, and TIP3P water molecules were used as the solvent. System neutrality was achieved by adding Na^+ ions. Energy minimization was carried out using the steepest descent algorithm. Equilibration was conducted in two stages: NVT (constant number of particles, volume, and temperature) and NPT (constant number of particles, pressure, and temperature) ensembles, each for 100,000 steps with a coupling constant of 0.1 ps and a total duration of 100 ps. Subsequently, a production run of free MD simulation was performed for 5,000,000 steps with a time

step of 2 fs, resulting in a simulation time of 100 ns. Post-simulation analysis included trajectory evaluations using built-in tools, calculating root-mean-square deviation (RMSD), root-mean-square fluctuation (RMSF), radius of gyration (Rg), and Solvent-Accessible Surface Area (SASA).

2.12. In Silico Trial Immune Simulation

The immune simulation component used the C-ImmSim web Server [55]. The simulation was performed in 270 steps, which is equivalent to 90 days in the real world. The default option was used for all HLA types of the host. Three experimental conditions were simulated: virus alone, vaccine alone, and vaccine followed by virus. The common sequence of the three proteins was used as a substitute for the viral antigen sequence.

3. Results

3.1. Epitopes Prediction and Selection

A total 16,848 sequence of HA, 15,283 sequence of NA, and 11,800 sequence of M1 were obtained in the NIH protein database according to the set restrictions.

After the alignment process, common sequences were generated for each of the three proteins. The sequence length of HA is 585 aa, NA is 467 aa, and M1 is 248 aa.

By NetMHCPan prediction, we found 45, 65, and 34 epitopes from HA, NA, and M1 that have a binding ability to the given MHC class I molecules (Table S2). By IEDB prediction, we found 56, 28, and 15 epitopes from HA, NA, and M1 that have a binding ability to the given MHC class II molecules (Table S3).

The B-cell epitopes predicted by three different methods were de-duplicated, resulting in 33, 38, and 26 epitopes for HA, NA, and M1, respectively. The scoring of epitopes by the three methods is shown in Table S4. Since different methods require different numbers of amino acids for epitopes, the length of these epitopes varies from 5 aa to 33 aa. The prediction results are shown in the Table S5.

To identify the epitopes with a better performance for vaccine design, all 340 epitopes were examined in three aspects. A total of 56 epitopes met the requirements for antigenicity, allergenicity, and toxicity. We further filtered the epitopes out of the idea of controlling the molecular weight of the recombinant protein and enhancing antigenicity. Using the VaxiJen model score of 0.78 as the new criterion, 24 epitopes that meet the new criterion can be obtained. The higher the VaxiJen model score, the higher the immunogenicity of epitopes, which is 0.78 to the optimal number of epitopes. Among these 24 epitopes, only one B-cell epitope of HA was found. Six additional B-cell epitopes of HA were selected under relaxed requirements for allergen testing, resulting in a total of 30 candidate epitopes, as shown in Table 1.

The epitope prediction was performed directly using the common sequence in the previous stage without considering the non-conserved regions in the sequence. This section examined the conservation of candidate epitopes in all antigen sequences. Three of the thirty candidate epitopes had less than 50% protein sequence matches with 100% homology and were eliminated. Twenty-one of the remaining epitopes matched over 95% of the protein sequences with 100% homology, indicating that the candidate epitopes are highly conserved, as shown in Table 1.

By calculating the distribution of corresponding Class I and II MHC molecules in the population, it can be determined that the candidate epitopes can induce immunity in 90.14% of the global population (Figure S2). The minimum number of epitope/HLA combination hits recognized by 90% of the population was 1.01. The average number of epitope hits/HLA combinations recognized by the population was 2.51.

Table 1. Physicochemical properties of the peptide.

ID	Sequence	Vaxijen	Toxinpred Result	AllerTOP	AllergenFP	Percent of Protein Sequence Matches at Identity ≥ 100%
>HA-B-7	^{355}GTKYRPP^{349}	0.7931	Non-Toxin	Non-Allergen	Non-Allergen	99.89%
>HA-B-ES-11	^{485}LERKLK^{480}	1.29	Non-Toxin	Allergen	Non-Allergen	99.79%
>HA-B-KT-11	^{313}EADCLHE^{307}	0.8468	Non-Toxin	Non-Allergen	Allergen	99.96%
>HA-B-KT-15	^{481}HLLALE^{476}	1.014	Non-Toxin	Non-Allergen	Allergen	99.91%
>HA-B-KT-3	^{48}VTGVIPLT^{41}	0.9925	Non-Toxin	Allergen	Non-Allergen	98.57%
>HA-B-KT-4	^{91}KLCPDCLNCTDLDVALGRPMCVGTT^{67}	1.4043	Non-Toxin	Allergen	Non-Allergen	48.72%
>HA-B-KT-5	^{112}KASILHEVRPVTSGCFPI^{95}	0.9508	Non-Toxin	Allergen	Non-Allergen	51.10%
>HA-T1-13	^{110}EVRPVTSGCF^{101}	1.4093	Non-Toxin	Non-Allergen	Non-Allergen	96.67%
>HA-T1-16	^{305}EADCLHEKY^{307}	1.3296	Non-Toxin	Non-Allergen	Non-Allergen	62.32%
>HA-T1-39	^{53}VIPLTTTPTK^{44}	0.9816	Non-Toxin	Non-Allergen	Non-Allergen	97.30%
>HA-T2-43	^{142}GYEHIRLSTQNVIDA^{128}	0.8272	Non-Toxin	Non-Allergen	Non-Allergen	0.00%
>HA-T2-55	^{401}TSHGAHGVAVAADLK^{387}	0.9048	Non-Toxin	Non-Allergen	Non-Allergen	99.20%
>M1-B-3	^{38}HCWFGGKEFD^{29}	1.7569	Non-Toxin	Non-Allergen	Non-Allergen	99.83%
>M1-B-4	^{110}RRCVSFH^{104}	1.2307	Non-Toxin	Non-Allergen	Non-Allergen	51.90%
>M1-B-6	^{94}LSGMGTTATKKK^{83}	1.2071	Non-Toxin	Non-Allergen	Non-Allergen	99.76%
>M1-B-9	^{143}NPGNYSMQVK^{134}	1.5183	Non-Toxin	Non-Allergen	Non-Allergen	99.55%
>M1-T1-1	^{183}RREMQMVSAM^{174}	1.0599	Non-Toxin	Non-Allergen	Non-Allergen	99.83%
>M1-T1-34	^{93}GMGTTATKK^{85}	1.4585	Non-Toxin	Non-Allergen	Non-Allergen	99.81%
>M1-T2-12	^{143}MVMYLNPGNYSMQVK^{129}	1.3408	Non-Toxin	Non-Allergen	Non-Allergen	99.42%
>M1-T2-13	^{46}FGGKEFDLDSALEWI^{32}	0.9846	Non-Toxin	Non-Allergen	Non-Allergen	99.69%
>M1-T2-14	^{140}YCLMVMYLNPGNYSM^{126}	1.276	Non-Toxin	Non-Allergen	Non-Allergen	99.31%
>M1-T2-3	^{106}KKKGLILAERKMRRC^{92}	1.3097	Non-Toxin	Non-Allergen	Non-Allergen	51.69%
>M1-T2-4	^{105}TKKKGLILAERKMRR^{91}	1.1846	Non-Toxin	Non-Allergen	Non-Allergen	51.69%
>NA-B-3	^{91}RLSCPGS^{85}	1.0799	Non-Toxin	Non-Allergen	Non-Allergen	98.60%
>NA-B-7	^{163}RHLISVKLG^{155}	1.2533	Non-Toxin	Non-Allergen	Non-Allergen	98.77%
>NA-T1-14	^{161}KLRHLISVK^{153}	1.1364	Non-Toxin	Non-Allergen	Non-Allergen	98.82%
>NA-T1-15	^{164}RHLISVKLGK^{155}	0.9099	Non-Toxin	Non-Allergen	Non-Allergen	98.70%
>NA-T1-23	^{114}ETKGNSAPL^{106}	0.9269	Non-Toxin	Non-Allergen	Non-Allergen	90.89%
>NA-T2-11	^{165}RNKLRHLISVKLGKI^{151}	0.8665	Non-Toxin	Non-Allergen	Non-Allergen	98.38%
>NA-T2-18	^{163}EDRNKLRHLISVKLG^{149}	0.97	Non-Toxin	Non-Allergen	Non-Allergen	49.06%

The numbers at the beginning and end of the sequence represent its start and end positions in the protein. ToxinPred result, AllerTOP, and AllergenFP are qualitative data, used to determine whether the sequence is toxic or allergenic.

3.2. Recombinant Multi-Epitope Vaccine

The selection yielded 27 epitopes that met the criteria. However, the sequences of seven of these epitopes were contained intact in other epitopes, and two M1 T-cell epitopes had the same core as other T-cell epitopes. After removing these duplicates there were 18 epitopes for the vaccine design. Two adjuvants human β-defensin 3 and LL37 were added to the vaccine constructs.

The N-terminal of the recombinant vaccine was two adjuvants, connected with an EAAAK linker. The adjuvant ends with 18 epitopes attached by EAAAK linkers. Several attempts were made to determine the final scheme for the arrangement of the subsequent epitopes in order to make the constructs fold well. The end of the adjuvant was connected to the cytotoxic T-lymphocytes (CTL) epitopes, and the GPGPG linkers were located between and after the CTL epitopes. GPGPG is a glycine-rich linker that provides free activity and a high degree of accessibility and flexibility to adjacent regions, in addition to increasing structural solubility [56]. Subsequently, the helper-T-lymphocytes (HTL) epitope was connected by KK linkers and the B-cell epitope was connected by AAY linkers. The sequence of the constructed vaccine contains 345 amino acids (Figure 2A).

The vaccine scored 0.5979 in the VaxiJen test and was judged as an antigen. Both allergenicity tests indicated that it was not an allergen. Predictions based on machine learning, MERCI, and BLAST showed that the vaccine had a hybrid score of 0.2 and was determined to be non-toxic.

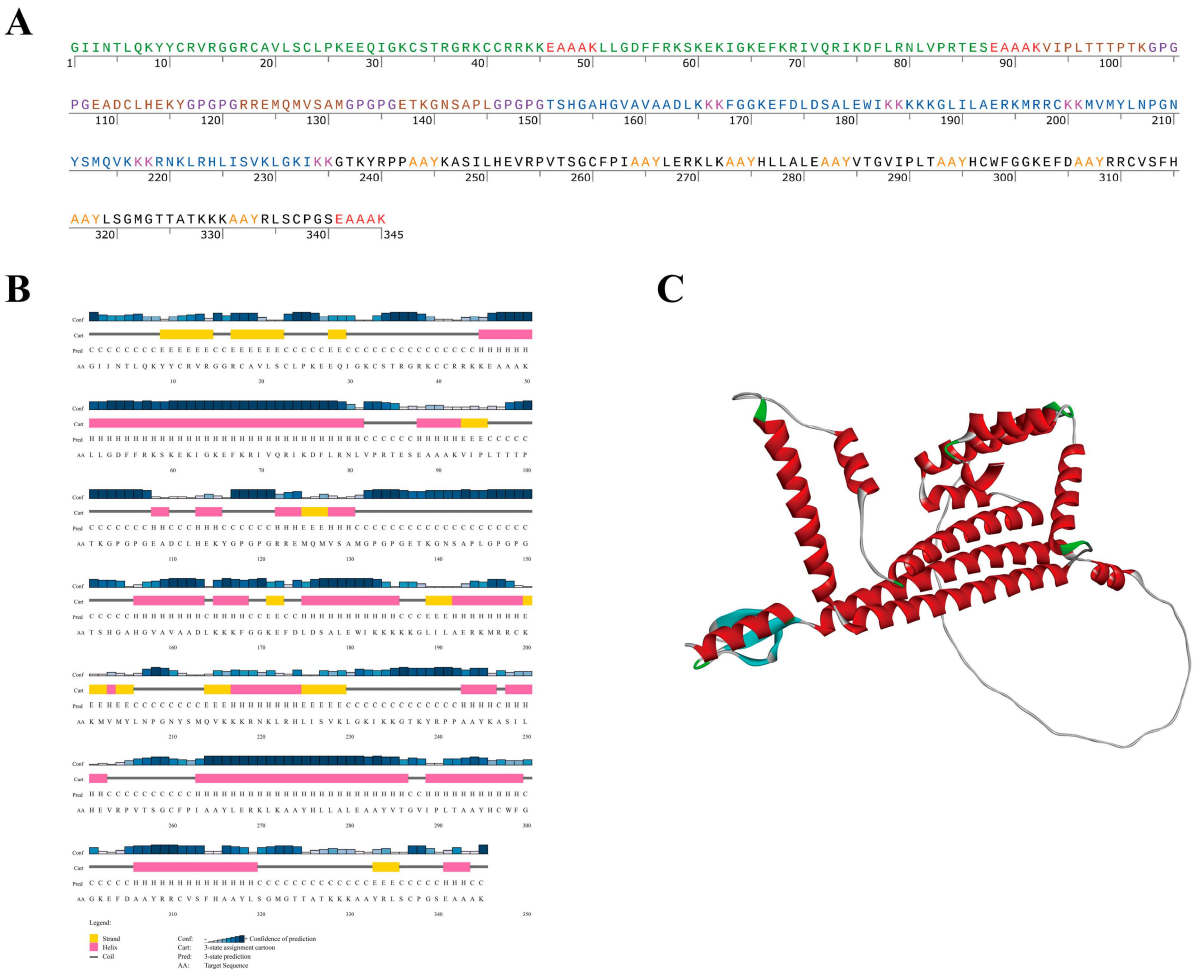

Figure 2. Sequence Alignment and Vaccine Structure (**A**): Protein sequence of the recombinant multi-epitope vaccine. The red, purple, pink, and orange sequences represent linkers, the green sequence represents the adjuvant peptide, and the other sequences represent the included epitopes. (**B**): Secondary structure prediction of the recombinant multi-epitope vaccine. (**C**): Tertiary structure prediction of the recombinant multi-epitope vaccine.

3.3. Properties of Multi-Epitope Vaccine

The predicted scaled solubility of the vaccine was 0.666. The population average of the experimental data set was 0.45, so any proportional solubility value greater than 0.45 was predicted to have a higher solubility than the average soluble *Escherichia coli* protein in the experimental solubility data set.

The recombinant vaccine consisted of 345 amino acids and had a molecular weight of 37.91 kDa. It had a theoretical pI of 10.08. The half-life of this protein was measured to be 30 h in mammalian reticulocytes, >20 h in yeast, and >10 h in *E. coli*. It was expected to reside in the body long enough to be effective. The instability index was 27.50. This index provides an estimate of the stability of the protein in tubes. Proteins with an instability index less than 40 are predicted to be stable and favorable for vaccine production and storage. The protein had an aliphatic index of 71.94, which also characterized its thermal stability. The GRAVY value of −0.46 indicated the hydrophilicity of the protein.

The secondary structure prediction revealed that the protein contained 43.48% coil, 44.64% helix, and 11.88% strand (Figure 2B).

The tertiary structure of the protein was predicted using AlphaFold2, and five different tertiary structures were generated. AlphaFold2 used the predicted local distance difference test loss (pLDDT) to directly derive per-position confidence for the predicted models. One

of these models was chosen as the tertiary structure of the protein based on the plausibility of the structure and pLDDT. The pLDDT score of this model was 66.45. This structure was subjected to repeated structural perturbation and subsequent overall structural relaxation by MD simulations using the Galaxy web server, and the optimized structure is shown in Figure 2C.

In the structure validation part, the designed recombinant protein passed all three tests set. Specifically, the overall quality factor of the recombinant protein was 99.64% in the ERRAT test (Figure 3A). The overall quality factor was expressed as the percentage of the protein for which the calculated error value falls below the 95% rejection limit. Good high-resolution structures generally produce values around 95% or higher. For lower resolutions the average overall quality factor is around 91%. ProCheck Ramachandran plots were made for the structures before and after optimization. The percentage of residues in the most favorable region before optimization was 79.1%, with 1.4% of residues in the disallowed region (Figure 3B). After optimization, the percentage of residues in the most favorable regions was 99.0%, and the disallowed regions had no residues (Figure 3C). Based on an analysis of 118 structures of a resolution of at least 2.0 Angstroms and an R-factor no greater than 20%, a good quality model would be expected to have over 90% in the most favored regions.

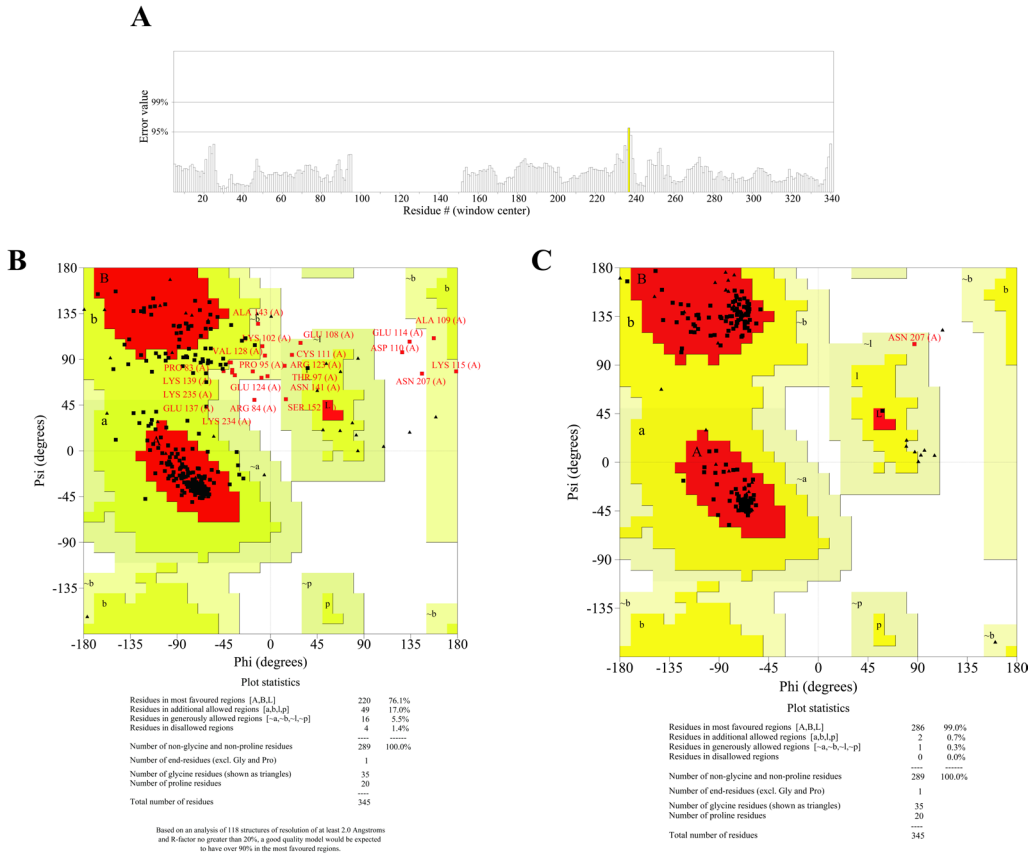

Figure 3. Structural Validation of the Vaccine (**A**): ERRAT test—the ERRAT test is used for overall structural validation of the vaccine. Structural regions that can be rejected at the 95% confidence level are yellow. Error values above this level are expected for 5% of good protein structures. Areas that can be rejected at the 99% level are shown in red. (**B**): Initial structure validation by ProCheck Ramachandran plots. ProCheck Ramachandran plot shows the distribution of dihedral angles φ (phi) and ψ (psi) of the vaccine structure. The most favorable regions are shown, with the red and yellow regions representing the most favorable angles. The quality of the structure is indicated by the number of residues in these regions, with marked outliers showing potential structural problem regions. (**C**): Refined structure validation by ProCheck Ramachandran plots.

3.4. Vaccine Performance Validation

The RMSD plot (Figure 4A) demonstrates the overall stability and conformational changes of the TLR4–vaccine complex compared to TLR4 and the vaccine alone. In the first ~5 ns, all systems showed initial fluctuations, which then stabilized, indicating that the system reached equilibrium under MD conditions. The TLR4–vaccine complex (black) exhibited less fluctuation than the vaccine alone (red), suggesting that the binding to TLR4 stabilizes the vaccine's conformation. This behavior indicates that the binding of TLR4 impacts the structural stability of the vaccine, but after 5 ns, the complex reached a relatively stable conformation, with an average RMSD value of 0.998 nm, confirming the equilibrium state of the complex.

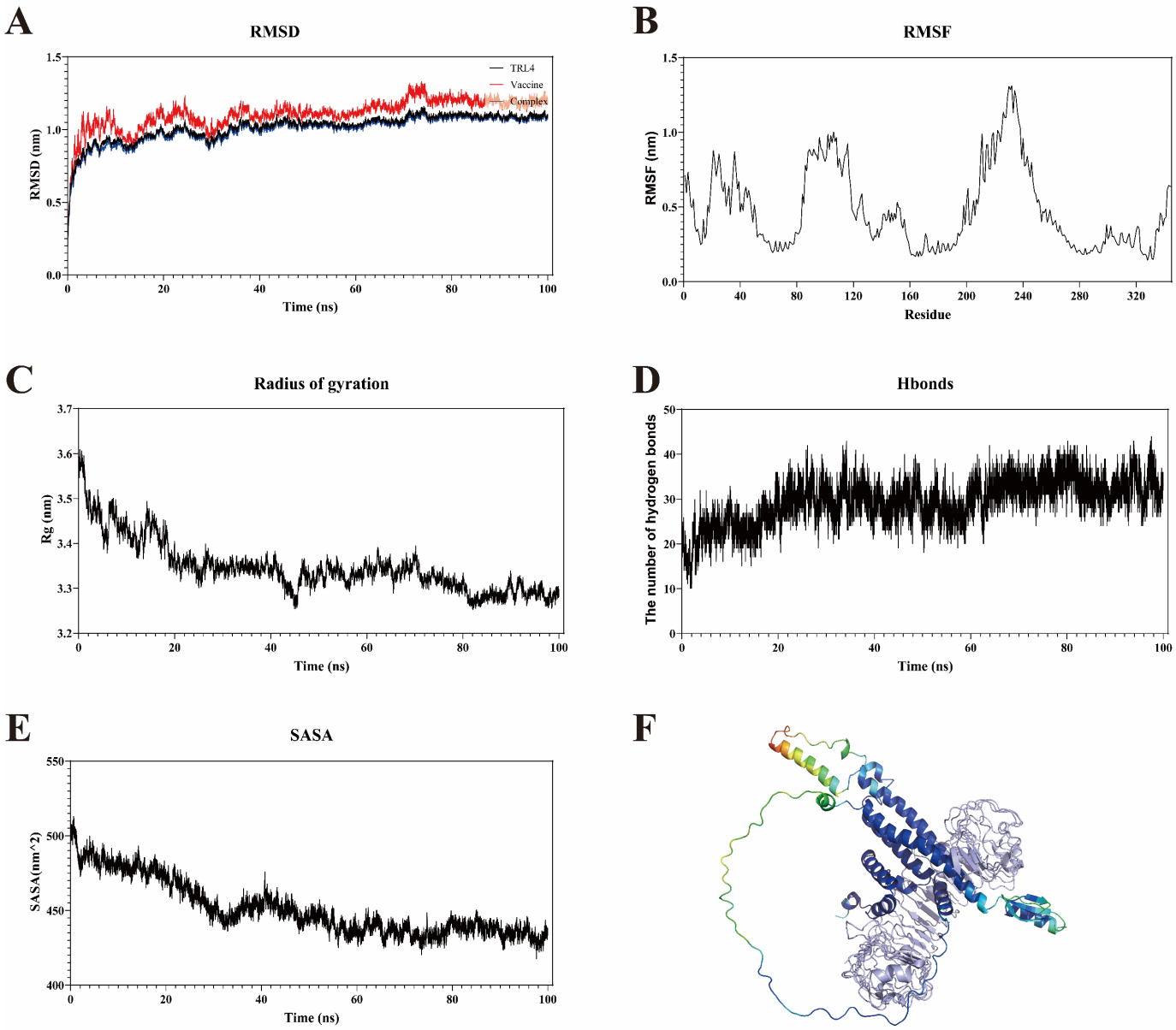

Figure 4. Molecular Dynamics Results. (**A**): RMSD Analysis. Black represents the RMSD of TLR4 during the simulation, red represents the RMSD of the vaccine during the simulation, and blue represents the RMSD of the complex during the simulation. (**B**): RMSF Analysis. (**C**): Radius of Gyration (Rg) Analysis. (**D**): Hydrogen Bond Analysis. (**E**): Solvent-Accessible Surface Area (SASA) Analysis. (**F**): A schematic diagram of the complex using B-factor as the color basis. Light blue represents TLR4, while dark blue to red represents the vaccine structure, with red indicating regions of high structural variability and blue indicating stable regions.

The RMSF plot (Figure 4B) reflects the flexibility of residues within the vaccine. Peaks around residues 30, 100, and 230 suggest regions with higher flexibility, corresponding to flexible loops observed in our tertiary structure prediction. Lower RMSF values in other regions indicate more stable areas of the vaccine, maintaining structural integrity and stable binding within the complex.

The Rg plot (Figure 4C) evaluates the compactness of the TLR4–vaccine complex. The stable Rg values indicate that the complex maintains a consistent tertiary structure without significant expansion or contraction, further supporting its structural stability. Minor fluctuations observed during the MD simulation are expected and do not indicate any significant unfolding or aggregation, reaffirming the compact nature of the complex.

The hydrogen bond analysis (Figure 4D) shows dynamic changes in the hydrogen bonds between TLR4 and the vaccine throughout the simulation, with an average of 29 dynamic hydrogen bonds. The fluctuations in the number of hydrogen bonds indicate transient but sustained interactions between TLR4 and the vaccine, crucial for the stability and specificity of the binding. These interactions are essential for maintaining the integrity of the complex and potentially influencing the binding affinity.

The SASA plot (Figure 4E) records the solvent accessible surface area of the TLR4–vaccine complex. The observed stability in SASA values, with minor fluctuations, suggests that the solvent exposure of the complex remains relatively constant. This consistency in SASA implies that the binding of the vaccine does not significantly alter the solvent-exposed regions of TLR4, which could be linked to its functionality and stability in the bound state.

The results (Figure 4F) indicate that the core binding region of the complex exhibits minimal structural fluctuations during the simulation period, while regions with higher variability are located farther from the core binding site. This observation underscores the stability of the complex's binding interactions.

Simulated immunization uses a non-experimental approach to validate the stimulation of the body's immune system by vaccine. The predicted results are shown in Figure 5.

In terms of antibody levels, the first thing that could be found was the ability of the vaccine to induce the production of all types of antibodies. Natural infection took about ten days or more to start producing various types of antibodies and reached its peak in the next ten days. In contrast, when infected with the virus after vaccination, all types of antibodies could be produced within a short period of time, and the antibody level could reach about half of the peak after ten days. At the same time, the rapid production of antibodies made it possible to inhibit the reproduction of the virus in a short time. A natural infection required up to twenty days to eliminate the pathogen, whereas after vaccination, this time was reduced to one or two days. The population of B cells in the body should explain the differences in the antibody levels. A response period of about ten days was required in a natural infection for B cells to initiate proliferation. After vaccination, however, the body carried a level of memory B cells that allowed it to undergo rapid proliferation.

The colony of both CTL and HTL cells explained, on the other hand, the role of vaccines in IBV infection. For HTL, the vaccination kept it at a high level for 40 to 50 days. Infection with the virus on day 30 stimulated HTL, which were in the late plateau phase to proliferate at a faster rate. In contrast, direct infection without vaccination had a response period of nearly ten days with no significant change in the HTL colony. Since the vaccine is a peptide, it is difficult to directly stimulate the proliferation of CTL cells. Its colony did not change significantly after vaccination. However, it could also be found that the organism invaded by the virus after direct vaccination could proliferate CTL rapidly.



Figure 5. Result of in silico trial immune simulation. (**A**): Immune Responses Induced by a Single Vaccination. (**B**): Immune Responses Induced by Viral Challenge Alone. (**C**): Immune Responses Induced by Vaccination Followed by Viral Challenge. From top to bottom, the induction of various antibody types, B cell proliferation, TH cell proliferation, TC cell proliferation, and cytokine induction are described, respectively.

Various cytokines had a similar profile to the antibodies. The vaccine itself could stimulate the synthesis of cytokines, while the infection after vaccination shortened the response period for cytokine production, allowing the rapid synthesis of cytokines to peak within one to two days.

Considering that T-cell epitopes need to be processed before being presented, we used separate T-cell epitopes instead of vaccine proteins for molecular docking. Firstly, the prediction of the epitope tertiary structure was still performed by AlphaFold2. Afterwards, two MHC class I molecules and four MHC class II molecules were selected for docking depending on the MHC molecule targeted by the epitope. The results of the docking were presented as RDOCK scores, listed in Table 2. As shown in Figure S3, the red helical structure is the antigen-binding groove of the MHC molecule, and the docked ligands (T-cell epitopes) are all in this domain. The RDOCK scores were all negative and most of them were less than −10, predicting a good binding activity. Binding to MHC molecules is necessary to stimulate cellular immunity.

Table 2. Docking result (RDOCK score).

	HLA-A*03:01	HLA-B*27:05	DRB1*0101	DRB3*0101	HLA-DPA1*02:01	HLA-DQA1*05:01
>HA-T1-13	−9.19282	−11.9748	na.	na.	na.	na.
>HA-T1-39	−12.7715	−20.2849	na.	na.	na.	na.
>HA-T1-16	−15.521	−16.2007	na.	na.	na.	na.
>M1-T1-34	−17.143	−19.8692	na.	na.	na.	na.
>M1-T1-1	−1.80865	−10.4786	na.	na.	na.	na.
>NA-T1-23	−10.6774	−11.5534	na.	na.	na.	na.
>NA-T1-14	−15.3003	−21.7556	na.	na.	na.	na.
>NA-T1-23	−13.7475	−12.7391	na.	na.	na.	na.
>HA-T2-55	na.	na.	−15.3999	−16.2973	−23.6671	−25.986
>M1-T2-13	na.	na.	−10.8606	−19.0344	−16.9406	−17.4544
>M1-T2-3	na.	na.	−14.7555	−5.11946	−23.8406	−13.2507
>M1-T2-12	na.	na.	−20.3404	−15.7837	−22.8958	−19.6213
>NA-T2-11	na.	na.	−17.1337	−10.9983	−21.1971	−10.525

4. Discussion

For a disease that can resemble the common cold, influenza packs a powerful and sometimes deadly punch. As many as 500,000 people die from influenza each year worldwide [57]. Influenza virus outbreaks are a constant reminder of the importance of developing a universal influenza vaccine. Yet conventional influenza vaccines lack the ability to provide comprehensive protection against different strains of the virus [58]. This is mainly due to two reasons. On the one hand, the antigenic variability of influenza viruses is significant [59]. In order to continue to spread in the population, influenza viruses must evolve through antigenic transfer and antigenic drift to evade herd immunity induced by natural infection and vaccination. On the other hand, it is because the technology used in conventional influenza vaccines does not have the ability to protect across subtypes. Widely used inactivated and live attenuated vaccines face inherent design limitations that hinder their ability to recognize highly variable viruses. In this research, we adopted a multi-epitope vaccine design protocol and selected IBV proteins from different lineages for vaccine design, which allowed us to obtain a vaccine with cross-subtype protection [58].

CTL and HTL epitopes were selected using the NetMHCpan and NetMHCIIpan servers, respectively, which predict peptide binding to MHC class I and II molecules using artificial neural networks. Both servers utilize a reference panel of 27 alleles from the IEDB, chosen for their representation of the most common specificities in the population and shared binding supertypes, providing > 97% coverage for class I and >99% for class II. The

predictions of the epitopes were all made using techniques that are highly accurate and relatively new. Since there were different techniques for predicting B-cell epitopes, we chose three representative methods. B-cell epitopes were predicted based on the surface accessible regions of the proteins [36], the chemical nature of amino acid residues [37], and sequence characteristics [38], respectively. In terms of screening conditions, broadly neutralizing antibodies against influenza virus HA epitopes has the potential to provide near-universal protection against influenza virus infection [60]. We relaxed the conditions on the screening of B-cell epitopes against HA. The criteria for antigenicity, toxicity, and conservation remained stringent, while allergenicity was permitted to fail in one test. Such a screening strategy allows for a more comprehensive and rigorous selection of epitopes that meet the required criteria.

Embedding the adjuvant sequence into the vaccine sequence is crucial because it enhances the vaccine's immunogenicity, promoting a stronger immune response. The adjuvant activates different components of the immune system, such as dendritic cells and macrophages, helping to boost antigen presentation and cytokine release, which in turn increases the reactivity of T cells and B cells. This strategy is particularly important for multi-epitope vaccines, which may have lower immunogenicity, as it helps to generate a stronger memory immune response, thereby improving the vaccine's efficacy and protective potential. In the design of the recombinant multi-epitope vaccine, we considered the effect of epitope ordering on the protein tertiary structure and on epitope recognition. So, we placed the adjuvant and B-cell epitopes at opposite ends of the sequence to increase the likelihood of their appearance on the surface and thus facilitate their recognition by the immune system. The order within the same class of epitopes was also examined to obtain better folded proteins. When analyzing the predicted secondary and tertiary structures, we could clearly see that they had many similarities and could corroborate each other. In this research, AlphaFold2 was chosen for the prediction of tertiary structures. There are two main reasons for this. The first reason is the surprising prediction accuracy when predicting the 3D coordinates of a protein given the amino acid sequence. Their model proved to be the most accurate of all submissions for more than 90% of the targets and, in most cases, fell quite short of the next best group [61,62]. It can be said that it is the best prediction accuracy solution available today. In addition, we also tried many other prediction strategies, such as I-TASSER, SWISS-MODEL, RaptorX, etc. The biggest problem was that most of these prediction strategies were based on homology modeling, which means that similar sequences were needed as templates. However, the vaccine sequence was designed to combine multiple fragments of three proteins, making it difficult to find sequences with high similarity. Ultimately, we chose the tertiary structure predicted by AlphaFold2. To improve the accuracy of the prediction structure, we used the Galaxy server to make improvements. By comparing several optimized structures before and after, we found that the Galaxy server was able to fix possible errors in the structure and improve the performance in the validation.

The MD simulation results demonstrated the stability of the TLR4–vaccine complex, highlighting the vaccine's potential for effective recognition by the innate immune system. For molecular docking, we focused on the T-cell epitope presentation process, using the constructed T-cell epitopes from the vaccine as ligands instead of the entire vaccine protein. The RDOCK-optimized conformations exhibited strong binding activity, supporting the vaccine's effectiveness in cellular immunity.

Simulated immunity results clearly showed the vaccine's protective effect against influenza virus challenges. However, it should be noted that the viral proteins used in this study were aligned to consensus sequences, and the vaccine epitopes were derived from these sequences. While this parameter setup allowed us to evaluate the vaccine's

performance against viruses with common characteristics, consensus sequences lack strain-specific features. Thus, the vaccine's efficacy against all IBV subtypes used in its design cannot be fully ascertained. The potential breadth of protection can only be inferred from the conservation of the vaccine epitopes, with most epitopes showing over 99% conservation across all analyzed proteins.

5. Conclusions

In this research, a recombinant multi-epitope vaccine against IBV was constructed. It had an excellent performance in antigenicity, allergenicity, toxicity, and conservancy tests. Population coverage studies showed that it triggered an immune response in 90.14% of the population worldwide. This research also predicted the secondary and tertiary structures of the recombinant proteins and verified the protective effect using both molecular docking, molecular dynamics, and simulated immunization methods.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/pr13030681/s1, Figure S1: Results of B-cell epitope prediction; Figure S2: Population coverage rate; Figure S3: Molecular docking; Table S1: Alleles for MHC; Table S2: Epitope prediction of MHC I; Table S3: Epitope prediction of MHC II; Table S4: B-cell epitope prediction; Table S5: Physicochemical properties of epitopes.

Author Contributions: Conceptualization, W.H. and Y.Y.; methodology, H.W.; data curation, Z.C.; writing—original draft preparation, H.W.; writing—review and editing, C.Z.; project administration, W.H. and H.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by The National Key Research and Development Program of China, grant number 2022YFC2304302.

Data Availability Statement: The original contributions presented in this study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding authors.

Conflicts of Interest: Author Hao Wu was employed by the company The Second Laboratory of Viral Vaccine Research, Wuhan Institute of Biological Products Co., Ltd., The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript.

influenza B viruses	IBV
Victoria lineage	BV
Yamagata lineage	BY
hemagglutinin	HA
matrix protein 1	M1
matrix protein 2	M2
The World Health Organization	WHO
National Center for Biotechnology Information	NCBI
Immune Epitope Database	IEDB
grand average of hydropathicity	GRAVY
root-mean-square deviation	RMSD
root-mean-square fluctuation	RMSF
cytotoxic T-lymphocytes	CTL
helper-T-lymphocytes	HTL
predicted local distance difference test loss	pLDDT
Radius of Gyration	Rg
Solvent-Accessible Surface Area	SASA

References

1. Iuliano, A.D.; Roguski, K.M.; Chang, H.H.; Muscatello, D.J.; Palekar, R.; Tempia, S.; Cohen, C.; Gran, J.M.; Schanzer, D.; Cowling, B.J.; et al. Estimates of global seasonal influenza-associated respiratory mortality: A modelling study. *Lancet* **2018**, *391*, 1285–1300. [CrossRef] [PubMed]
2. Yang, J.; Lau, Y.C.; Wu, P.; Feng, L.; Wang, X.; Chen, T.; Ali, S.T.; Peng, Z.; Fang, V.J.; Zhang, J.; et al. Variation in Influenza B Virus Epidemiology by Lineage, China. *Emerg. Infect. Dis.* **2018**, *24*, 1536–1540. [CrossRef] [PubMed]
3. Caini, S.; Kusznierz, G.; Garate, V.V.; Wangchuk, S.; Thapa, B.; de Paula Junior, F.J.; Ferreira de Almeida, W.A.; Njouom, R.; Fasce, R.A.; Bustos, P.; et al. The epidemiological signature of influenza B virus and its B/Victoria and B/Yamagata lineages in the 21st century. *PLoS ONE* **2019**, *14*, e0222381. [CrossRef] [PubMed]
4. Zaraket, H.; Hurt, A.C.; Clinch, B.; Barr, I.; Lee, N. Burden of influenza B virus infection and considerations for clinical management. *Antiviral Res.* **2021**, *185*, 104970. [CrossRef] [PubMed]
5. Shang, M.; Blanton, L.; Brammer, L.; Olsen, S.J.; Fry, A.M. Influenza-Associated Pediatric Deaths in the United States, 2010–2016. *Pediatrics* **2018**, *141*, e20172918. [CrossRef]
6. Jennings, L.; Huang, Q.S.; Barr, I.; Lee, P.I.; Kim, W.J.; Buchy, P.; Sanicas, M.; Mungall, B.A.; Chen, J. Literature review of the epidemiology of influenza B disease in 15 countries in the Asia-Pacific region. *Influenza Other Respir. Viruses* **2018**, *12*, 383–411. [CrossRef]
7. Paul Glezen, W.; Schmier, J.K.; Kuehn, C.M.; Ryan, K.J.; Oxford, J. The burden of influenza B: A structured literature review. *Am. J. Public Health* **2013**, *103*, e43–e51. [CrossRef]
8. Borchering, R.K.; Gunning, C.E.; Gokhale, D.V.; Weedop, K.B.; Saeidpour, A.; Brett, T.S.; Rohani, P. Anomalous influenza seasonality in the United States and the emergence of novel influenza B viruses. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2012327118. [CrossRef]
9. Rota, P.A.; Wallis, T.R.; Harmon, M.W.; Rota, J.S.; Kendal, A.P.; Nerome, K. Cocirculation of two distinct evolutionary lineages of influenza type B virus since 1983. *Virology* **1990**, *175*, 59–68. [CrossRef]
10. Virk, R.K.; Jayakumar, J.; Mendenhall, I.H.; Moorthy, M.; Lam, P.; Linster, M.; Lim, J.; Lin, C.; Oon, L.L.E.; Lee, H.K.; et al. Divergent evolutionary trajectories of influenza B viruses underlie their contemporaneous epidemic activity. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 619–628. [CrossRef]
11. Caini, S.; Meijer, A.; Nunes, M.C.; Henaff, L.; Zounon, M.; Boudewijns, B.; Del Riccio, M.; Paget, J. Probable extinction of influenza B/Yamagata and its public health implications: A systematic literature review and assessment of global surveillance databases. *Lancet Microbe* **2024**, *5*, 100851. [CrossRef] [PubMed]
12. Reich, S.; Guilligay, D.; Pflug, A.; Malet, H.; Berger, I.; Crepin, T.; Hart, D.; Lunardi, T.; Nanao, M.; Ruigrok, R.W.; et al. Structural insight into cap-snatching and RNA synthesis by influenza polymerase. *Nature* **2014**, *516*, 361–366. [CrossRef] [PubMed]
13. Li, L.; Chai, Y.; Peng, W.; Li, D.; Sun, L.; Gao, G.F.; Qi, J.; Xiao, H.; Liu, W.J.; von Itzstein, M.; et al. Structural and inhibitor sensitivity analysis of influenza B-like viral neuraminidases derived from Asiatic toad and spiny eel. *Proc. Natl. Acad. Sci. USA* **2022**, *119*, e2210724119. [CrossRef] [PubMed]
14. Wilson, I.A.; Skehel, J.J.; Wiley, D.C. Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 A resolution. *Nature* **1981**, *289*, 366–373. [CrossRef] [PubMed]
15. Harrison, S.C. Viral membrane fusion. *Virology* **2015**, *479–480*, 498–507. [CrossRef]
16. Myers, M.L.; Gallagher, J.R.; Woolfork, D.D.; Stradtmann-Carvalho, R.K.; Maldonado-Puga, S.; Bock, K.W.; Boyoglu-Barnum, S.; Syeda, H.; Creanga, A.; Alves, D.A.; et al. Impact of adjuvant: Trivalent vaccine with quadrivalent-like protection against heterologous Yamagata-lineage influenza B virus. *Front. Immunol.* **2022**, *13*, 1002286. [CrossRef]
17. Rappuoli, R.; Bottomley, M.J.; D'Oro, U.; Finco, O.; De Gregorio, E. Reverse vaccinology 2.0: Human immunology instructs vaccine antigen design. *J. Exp. Med.* **2016**, *213*, 469–481. [CrossRef]
18. Creytens, S.; Pascha, M.N.; Ballegeer, M.; Saelens, X.; de Haan, C.A.M. Influenza Neuraminidase Characteristics and Potential as a Vaccine Target. *Front. Immunol.* **2021**, *12*, 786617. [CrossRef]
19. Chen, Y.Q.; Wohlbold, T.J.; Zheng, N.Y.; Huang, M.; Huang, Y.; Neu, K.E.; Lee, J.; Wan, H.; Rojas, K.T.; Kirkpatrick, E.; et al. Influenza Infection in Humans Induces Broadly Cross-Reactive and Protective Neuraminidase-Reactive Antibodies. *Cell* **2018**, *173*, 417–429.e10. [CrossRef]
20. Selzer, L.; Su, Z.; Pintilie, G.D.; Chiu, W.; Kirkegaard, K. Full-length three-dimensional structure of the influenza A virus M1 protein and its organization into a matrix layer. *PLoS Biol.* **2020**, *18*, e3000827. [CrossRef]
21. World Health Organization. *WHO Preferred Product Characteristics for Next Generation Influenza Vaccines*; World Health Organization: Geneva, Switzerland, 2017.
22. Tsybalova, L.M.; Stepanova, L.A.; Ramsay, E.S.; Vasin, A.V. Influenza B: Prospects for the Development of Cross-Protective Vaccines. *Viruses* **2022**, *14*, 1323. [CrossRef] [PubMed]

23. Grohskopf, L.A.; Blanton, L.H.; Ferdinands, J.M.; Chung, J.R.; Broder, K.R.; Talbot, H.K.; Morgan, R.L.; Fry, A.M. Prevention and Control of Seasonal Influenza with Vaccines: Recommendations of the Advisory Committee on Immunization Practices—United States, 2022–2023 Influenza Season. *MMWR Recomm. Rep.* **2022**, *71*, 1–28. [CrossRef]
24. Cardenas-Garcia, S.; Caceres, C.J.; Rajao, D.; Perez, D.R. Reverse genetics for influenza B viruses and recent advances in vaccine development. *Curr. Opin. Virol.* **2020**, *44*, 191–202. [CrossRef] [PubMed]
25. National Immunization Advisory Committee (NIAC) Technical Working Group (TWG), Influenza Vaccination TWG. [Technical guidelines for seasonal influenza vaccination in China (2021–2022)]. *Zhonghua Liu Xing Bing Xue Za Zhi* **2021**, *42*, 1722–1749. [CrossRef]
26. Ambrose, C.S.; Levin, M.J. The rationale for quadrivalent influenza vaccines. *Hum. Vaccin. Immunother.* **2012**, *8*, 81–88. [CrossRef] [PubMed]
27. Rahman, S.; Hasan, M.; Alam, M.S.; Uddin, K.M.M.; Moni, S.; Rahman, M. The evolutionary footprint of influenza A subtype H3N2 strains in Bangladesh: Implication of vaccine strain selection. *Sci. Rep.* **2022**, *12*, 16186. [CrossRef] [PubMed]
28. Skowronski, D.M.; Masaro, C.; Kwindt, T.L.; Mak, A.; Petric, M.; Li, Y.; Sebastian, R.; Chong, M.; Tam, T.; De Serres, G. Estimating vaccine effectiveness against laboratory-confirmed influenza using a sentinel physician network: Results from the 2005-2006 season of dual A and B vaccine mismatch in Canada. *Vaccine* **2007**, *25*, 2842–2851. [CrossRef]
29. Flannery, B.; Kondor, R.J.G.; Chung, J.R.; Gaglani, M.; Reis, M.; Zimmerman, R.K.; Nowalk, M.P.; Jackson, M.L.; Jackson, L.A.; Monto, A.S.; et al. Spread of Antigenically Drifted Influenza A(H3N2) Viruses and Vaccine Effectiveness in the United States During the 2018–2019 Season. *J. Infect. Dis.* **2020**, *221*, 8–15. [CrossRef]
30. Ren, H.; Zhou, P. Epitope-focused vaccine design against influenza A and B viruses. *Curr. Opin. Immunol.* **2016**, *42*, 83–90. [CrossRef]
31. Wang, W.C.; Sayedahmed, E.E.; Sambhara, S.; Mittal, S.K. Progress towards the Development of a Universal Influenza Vaccine. *Viruses* **2022**, *14*, 1684. [CrossRef]
32. Edgar, R.C. High-accuracy alignment ensembles enable unbiased assessments of sequence homology and phylogeny. *bioRxiv* **2022**. [CrossRef]
33. Waterhouse, A.M.; Procter, J.B.; Martin, D.M.; Clamp, M.; Barton, G.J. Jalview Version 2—A multiple sequence alignment editor and analysis workbench. *Bioinformatics* **2009**, *25*, 1189–1191. [CrossRef] [PubMed]
34. NetMHCpan—4.1. Available online: https://services.healthtech.dtu.dk/service.php?NetMHCpan-4.1 (accessed on 1 May 2023).
35. MHC-II Binding Predictions. Available online: http://tools.iedb.org/mhcii/ (accessed on 1 May 2023).
36. Emini, E.A.; Hughes, J.V.; Perlow, D.S.; Boger, J. Induction of hepatitis A virus-neutralizing antibody by a virus-specific synthetic peptide. *J. Virol.* **1985**, *55*, 836–839. [CrossRef] [PubMed]
37. Kolaskar, A.S.; Tongaonkar, P.C. A semi-empirical method for prediction of antigenic determinants on protein antigens. *FEBS Lett.* **1990**, *276*, 172–174. [CrossRef]
38. Jespersen, M.C.; Peters, B.; Nielsen, M.; Marcatili, P. BepiPred-2.0: Improving sequence-based B-cell epitope prediction using conformational epitopes. *Nucleic Acids Res.* **2017**, *45*, W24–W29. [CrossRef]
39. Doytchinova, I.A.; Flower, D.R. VaxiJen: A server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinform.* **2007**, *8*, 4. [CrossRef]
40. VaxiJen v2.0. Available online: http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html (accessed on 1 May 2023).
41. AllerTOP v2.0. Available online: https://www.ddg-pharmfac.net/AllerTOP/ (accessed on 1 May 2023).
42. AllergenFP v1.0. Available online: http://ddg-pharmfac.net/AllergenFP/ (accessed on 1 May 2023).
43. Gupta, S.; Kapoor, P.; Chaudhary, K.; Gautam, A.; Kumar, R.; Open Source Drug Discovery, C.; Raghava, G.P. In silico approach for predicting toxicity of peptides and proteins. *PLoS ONE* **2013**, *8*, e73957. [CrossRef]
44. ToxinPred. Available online: http://crdd.osdd.net/raghava/toxinpred/ (accessed on 1 May 2023).
45. Bui, H.H.; Sidney, J.; Li, W.; Fusseder, N.; Sette, A. Development of an epitope conservancy analysis tool to facilitate the design of epitope-based diagnostics and vaccines. *BMC Bioinform.* **2007**, *8*, 361. [CrossRef]
46. Epitope Conservancy Analysis. Available online: http://tools.iedb.org/conservancy/ (accessed on 1 May 2023).
47. Population Coverage Tool. Available online: http://tools.iedb.org/population/ (accessed on 1 May 2023).
48. Funderburg, N.; Lederman, M.M.; Feng, Z.; Drage, M.G.; Jadlowsky, J.; Harding, C.V.; Weinberg, A.; Sieg, S.F. Human β-defensin-3 activates professional antigen-presenting cells via Toll-like receptors 1 and 2. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 18631–18635. [CrossRef]
49. Keshri, A.K.; Rawat, S.S.; Chaudhary, A.; Sharma, S.; Kapoor, A.; Mehra, P.; Kaur, R.; Mishra, A.; Prasad, A. LL-37, the master Antimicrobial Peptide, its multifaceted role from combating infections to cancer immunity. *Int. J. Antimicrob. Agents* **2024**, *65*, 107398. [CrossRef]
50. Toxinpred2. Available online: https://webs.iiitd.edu.in/raghava/toxinpred2/ (accessed on 1 May 2023).
51. Protein-Sol. Available online: https://protein-sol.manchester.ac.uk/ (accessed on 1 May 2023).

52. Wilkins, M.R.; Gasteiger, E.; Bairoch, A.; Sanchez, J.C.; Williams, K.L.; Appel, R.D.; Hochstrasser, D.F. Protein identification and analysis tools in the ExPASy server. *Methods Mol. Biol.* **1999**, *112*, 531–552. [CrossRef] [PubMed]
53. ProtParam. Available online: https://web.expasy.org/protparam/ (accessed on 1 May 2023).
54. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Zidek, A.; Potapenko, A.; et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **2021**, *596*, 583–589. [CrossRef] [PubMed]
55. c-IMMSIM. Available online: https://150.146.2.1/C-IMMSIM/ (accessed on 1 May 2023).
56. Tarrahimofrad, H.; Rahimnahal, S.; Zamani, J.; Jahangirian, E.; Aminzadeh, S. Designing a multi-epitope vaccine to provoke the robust immune response against influenza A H7N9. *Sci. Rep.* **2021**, *11*, 24485. [CrossRef] [PubMed]
57. Brody, H. Influenza. *Nature* **2019**, *573*, S49. [CrossRef]
58. Nguyen, Q.T.; Choi, Y.K. Targeting Antigens for Universal Influenza Vaccine Development. *Viruses* **2021**, *13*, 973. [CrossRef]
59. Petrova, V.N.; Russell, C.A. The evolution of seasonal influenza viruses. *Nat. Rev. Microbiol.* **2018**, *16*, 47–60. [CrossRef]
60. Guthmiller, J.J.; Han, J.; Utset, H.A.; Li, L.; Lan, L.Y.; Henry, C.; Stamper, C.T.; McMahon, M.; O'Dell, G.; Fernandez-Quintero, M.L.; et al. Broadly neutralizing antibodies target a haemagglutinin anchor epitope. *Nature* **2022**, *602*, 314–320. [CrossRef]
61. Cramer, P. AlphaFold2 and the future of structural biology. *Nat. Struct. Mol. Biol.* **2021**, *28*, 704–705. [CrossRef]
62. Jones, D.T.; Thornton, J.M. The impact of AlphaFold2 one year on. *Nat. Methods* **2022**, *19*, 15–20. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.