

Article

Identification and In Silico Characterization of a Conserved Peptide on Influenza Hemagglutinin Protein: A New Potential Antigen for Universal Influenza Vaccine Development

Atin Khalaj-Hedayati ^{1,2,*} , Seyedehmaryam Moosavi ³ , Otilia Manta ^{4,5,6} , Mohamed H. Helal ⁷,
Mohamed M. Ibrahim ⁸, Zeinhom M. El-Bahy ⁹ and Ganden Supriyanto ^{1,*}

- ¹ Department of Chemistry, Faculty of Science and Technology, Airlangga University, Mulyorejo, Surabaya 60115, Indonesia
- ² School of Biosciences, Faculty of Health and Medical Sciences, Taylor's University, Subang Jaya 47500, Malaysia
- ³ Department of Nanotechnology Engineering, Faculty of Advance Technology and Multidiscipline, Airlangga University, Mulyorejo, Surabaya 60115, Indonesia; m.moosavi1987@gmail.com
- ⁴ Romanian Academy, Victor Slavescu Centre for Financial and Monetary Research, 050731 Bucharest, Romania; otilia.manta@rgic.ro
- ⁵ Romanian Academy, CE-MONT Mountain Economy Center, 725700 Vatra Dornei, Romania
- ⁶ Research Department, Romanian American University, 012101 Bucharest, Romania
- ⁷ Department of Chemistry, Faculty of Arts and Science, Northern Border University, Rafha 76413, Saudi Arabia; mohammed.hlal@nbu.edu.sa
- ⁸ Department of Chemistry, College of Science, Taif University, P.O. Box 11099, Taif 21944, Saudi Arabia; ibrahim@tu.edu.sa
- ⁹ Department of Chemistry, Faculty of Science, Al-Azhar University, Nasr City 11884, Egypt; zeinelbahy@azhar.edu.eg
- * Correspondence: atin.hedayati@gmail.com (A.K.-H.); ganden-s@fst.unair.ac.id (G.S.)



Citation: Khalaj-Hedayati, A.; Moosavi, S.; Manta, O.; Helal, M.H.; Ibrahim, M.M.; El-Bahy, Z.M.; Supriyanto, G. Identification and In Silico Characterization of a Conserved Peptide on Influenza Hemagglutinin Protein: A New Potential Antigen for Universal Influenza Vaccine Development. *Nanomaterials* **2023**, *13*, 2796. <https://doi.org/10.3390/nano13202796>

Academic Editor: Thomas Dippong

Received: 8 September 2023

Revised: 13 October 2023

Accepted: 17 October 2023

Published: 20 October 2023



Copyright: © 2023 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/>).

Abstract: Antigenic changes in surface proteins of the influenza virus may cause the emergence of new variants that necessitate the reformulation of influenza vaccines every year. Universal influenza vaccine that relies on conserved regions can potentially be effective against all strains regardless of any antigenic changes and as a result, it can bring enormous public health impact and economic benefit worldwide. Here, a conserved peptide (HA_{288–107}) on the stalk domain of hemagglutinin glycoprotein is identified among highly pathogenic influenza viruses. Five top-ranked B-cell and twelve T-cell epitopes were recognized by epitope mapping approaches and the corresponding Human Leukocyte Antigen alleles to T-cell epitopes showed high population coverage (>99%) worldwide. Moreover, molecular docking analysis indicated that *VLMENERTL* and *WTYNAELLV* epitopes have high binding affinity to the antigen-binding groove of the HLA-A*02:01 and HLA-A*68:02 molecules, respectively. Theoretical physicochemical properties of the peptide were assessed to ensure its thermostability and hydrophilicity. The results suggest that the HA_{288–107} peptide can be a promising antigen for universal influenza vaccine design. However, in vitro and in vivo analyses are needed to support and evaluate the effectiveness of the peptide as an immunogen for vaccine development.

Keywords: epitope mapping; immunoinformatic; hemagglutinin; nanoparticle; peptide-based vaccine; universal influenza vaccine

1. Introduction

Influenza is considered as one of the main human health problems in the world that causes respiratory diseases with 3 to 5 million cases of severe illness and up to 500,000 deaths reported every year [1]. Influenza A virus is classified into subtypes based on two surface antigens, neuraminidase (NA) and hemagglutinin (HA) glycoproteins. Up to now, 18 different HA and 11 different NA proteins have been identified [2]. The pathogenic

influenza A virus strains that are currently circulating in the human population include H1N1 and H3N2 and the strains that have a high probability for future pandemic outbreaks are H7N9, H9N2, H5N1, and H2N2 [3,4]. The most effective method to prevent influenza infection is vaccination and currently, most influenza vaccines focus on raising antibodies against the HA glycoprotein. This surface glycoprotein regulates the penetration of the virus into the host cells and it consists of a globular head (HA1) and a stalk region (HA2) [5,6].

Antigenic changes in surface proteins may cause the emergence of new variants of the influenza virus that necessitate the reformulation of influenza vaccines every year [7]. The lack of proof-reading activity of the viral polymerase enzyme is the main reason for changes in the genome, leading to variation in their surface antigens [8]. These antigenic changes allow the influenza viruses to escape host antibody immunity built up through previous vaccination or exposure and decrease the vaccine efficiency [9]. Hence, the circulating subtypes must be recognized annually to reformulate the influenza vaccine. The process relies on surveillance, genotype sequencing, and the measurement of the antigenic properties of the circulating strains. Sometimes the selected strains for vaccine composition may not match with the near future influenza viruses well enough, which may result in increased clinical cases and reduced vaccine protection [9,10]. Moreover, the emergence of new variants with distinct antigenic properties may happen due to the genetic reassortment that can bring up the pandemic with considerable illness and mortality [11]. Currently, the efficacy of the current influenza vaccines is at 44%, therefore, the National Institute of Allergy and Infectious Diseases recently recognized the development of an influenza vaccine with an efficacy of at least 75% as a high scientific priority in which epidemiological impacts of seasonal influenza would be reduced significantly [12]. Hence, the primary goal in influenza vaccine development is to direct the immune system to induce effective responses against multiple subtypes of the virus despite antigenic differences. The strategy to achieve this target is based on modern subunit vaccine development that uses conserved epitopes as vaccine composition to produce a vaccine with broad protection. The main aim is to decrease the occurrence, hospitalization, and death by at least 95%, which would also save \$3.5 billion each year in direct medical costs related to influenza. This economic benefit exceeds the proposed and current \$330 million in funding for the development of a universal influenza vaccine [13].

Universal vaccine design and development requires precise molecular and physicochemical knowledge of the antigen and its interactions with immune systems. In this regard, *in silico* tools can provide helpful information to facilitate antigen selection before vaccine development and laboratory experiments [14]. These technologies combine immunogenetics with computational tools to increase antigen selection accuracy based on physicochemical characteristics [15]. Antigen selection, B and T-cell epitope mapping, population coverage, and evaluation of physicochemical properties before conducting *in vitro* and *in vivo* experiments, can accelerate the development of the novel influenza vaccine. As studies showed that the antibodies produced against the HA2 domain have a higher neutralization breadth compared to HA1 antibodies due to the highest number of conserved epitopes, here, we evaluated HA2 amino acid sequences of six highly pathogenic influenza virus strains (H5N1, H7N9, H9N2, H1N1, H3N2, and H2N2) among human populations by utilizing computational approaches to identify conserved and antigenic peptide [16]. The set of results presented here contributes to the rational design of a universal influenza vaccine based on the conserved peptide that can potentially trigger broad immune responses.

2. Materials and Methods

The specific steps involved in identifying and characterizing the influenza conserved peptide are described in detail and a ketch of the entire workflow is demonstrated in Figure 1.

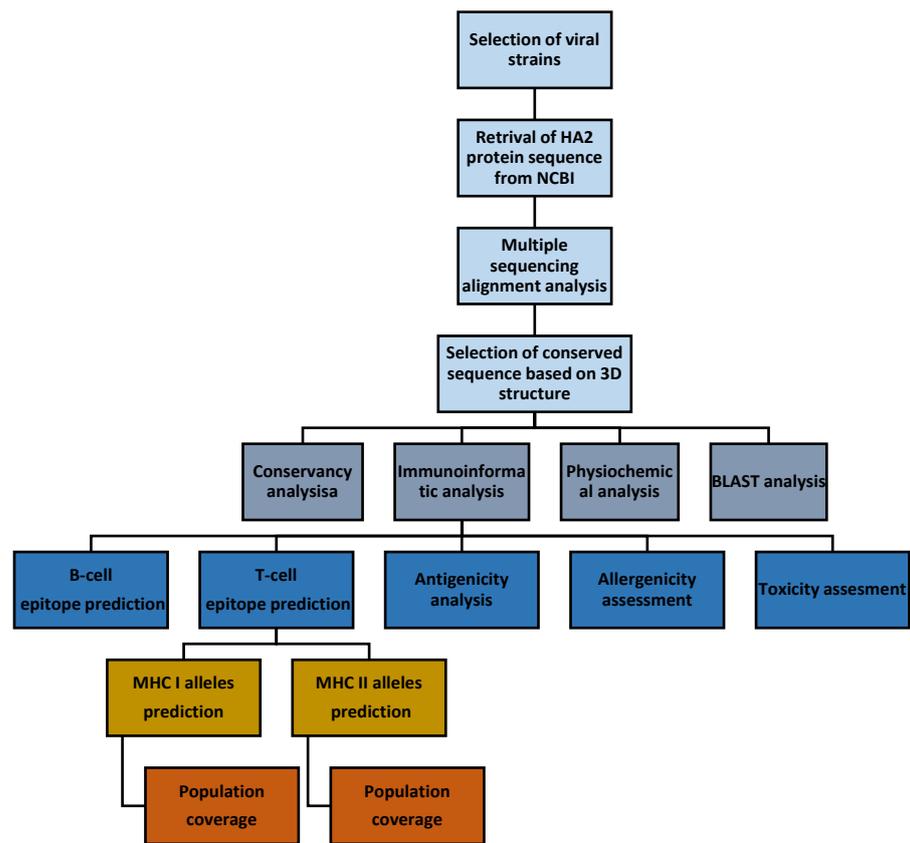


Figure 1. Workflow for in silico prediction and characterization analysis of a conserved peptide on HA glycoprotein as a potential antigen for universal influenza vaccine design.

2.1. Multiple Sequence Alignment, Structure, and Conservancy Analysis

A total of six highly pathogenic influenza virus strains (H5N1, H7N9, H9N2, H2N2, H1N1, and H3N2) were selected based on the Center for Disease Control and Prevention surveillance and report. The amino acid sequences of the HA2 domain of the selected strains were retrieved from the National Centre for Biotechnology Information (NCBI) database (<http://www.ncbi.nlm.nih.gov/protein>, access date: 1 July 2023) [3]. The retrieved amino acid sequences were subjected to multiple sequence alignment by Multalin software (<http://multalin.toulouse.inra.fr/multalin/>, access date: 1 July 2023) to obtain the conserved regions. Interactive 3D structure viewer, iCn3d, (<https://www.ncbi.nlm.nih.gov/Structure/icn3d/full.html>, access date: 1 July 2023) was used to select a conserved sequence on the HA protein based on the position of the sequence. The conservancy analysis tool from the Immune Epitope Database (IEDB) was used to estimate the identity percentage of the selected peptide in each strain [17]. Also, an NCBI Basic Local Alignment Search Tool (BLAST) analysis (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>, access date: 1 July 2023) was performed to confirm that the selected sequence is 100% influenza virus-related.

2.2. Identification of Linear B-Cell Epitopes

Hydrophilicity, surface accessibility, antigenicity, flexibility, and beta-turn of the candidate peptide were analyzed using different tools from IEDB (<http://tools.iedb.org/bcell/>, access date: 1 July 2023). The Emini method was used to predict linear B-cell epitope based on surface accessibility [18] and the Karplus and Schulz method was applied for flexibility prediction [19]. Chou and Fasman [20], Parker [21], and Kolaskar and Tongaonkar [22] methods were utilized to predict possible B-cell epitopes based on beta-turn, hydrophilicity, and antigenicity, respectively. Default threshold scores were used for all the methods and the residues with value above the threshold are considered to be part of an epitope.

2.3. Identification of T-Cell Epitopes and Population Coverage Analysis

The T-cell epitope prediction tool from IEDB was applied to calculate the binding affinity of the selected peptide to the most frequent Major Histocompatibility Complex (MHC) class I and II in the human population based on software recommendation (<http://tools.iedb.org/main/tcell/>, access date: 1 July 2023). For the best T-cell epitope selection among all predicted epitopes, the identified epitopes with percentile rank less or equal to 1% for MHC I and less and equal to 10% for MHC II were considered as the potential MHC binder epitopes [23,24]. Later, respective corresponding alleles to the MHC I and II epitopes on the influenza HA2 domain were evaluated for population coverage against the whole world human population using the IEDB population coverage analysis server (<http://tools.iedb.org/population/>, access date: 1 July 2023). The area option was set to “world” and the calculation option was set to “Class I and II combined”.

2.4. Cluster Analysis of the MHC-Restricted Alleles

The MHCcluster 2.0 server (<http://www.cbs.dtu.dk/services/MHCcluster/>, access date: 5 July 2023) was used to provide pictorial tree-based visualizations and highly instinctive heat map of the functional alliance between the MHC variants that predicted to interact with the selected peptide. The graphical tree and statistic heat map as analysis outputs describe the functional relationship between the MHC alleles from the overlap prediction binding specificity.

2.5. Molecular Docking Analysis

The MHC I alleles with the lowest percentile rank for each corresponding epitope (underlined alleles in Table 4) were subjected to docking analysis using the ClusPro 2.0 server (<https://cluspro.bu.edu/login.php>, access date: 5 July 2023). The 3D structure of MHC I proteins was retrieved from the Research Collaboratory for Structural Bioinformatics Protein Data Bank (PDB) and was refined by Autodock software (<http://autodock.scripps.edu>, access date: 5 July 2023) to prepare the pre-docking structure. Water molecules were removed and the supporting structure, beta-microglobulin, was retained due to their function of providing stability to the proteins in the host [25]. Also, the FASTA format of MHC I epitopes was converted into PDB format through OpenBabel GUI version 2.3. software, to analyze the interactions with the corresponding MHC molecules (<https://openbabel.org/docs/dev/index.html>, access date: 5 July 2023). The PDB files of both the receptors and the epitopes were uploaded to the ClusPro 2.0 server to acquire the desirable complexes in terms of free binding energy and better electrostatic interaction. The best models were selected based on the lower global binding energy. The MHC II alleles were not considered for the docking analysis as the structure of the predicted alleles is not available in the PDB database.

2.6. Antigenicity, Allergenicity, and Toxicity Assessment

The VaxiJen 2.0 server (<http://www.ddgpharmfac.net/vaxijen/>, access date: 5 July 2023) was applied to determine the antigenicity of the selected peptide with the threshold value of 0.4 (default) and the virus was selected as the target organism. The AllergenFP online server (<http://www.ddg-pharmfac.net/AllergenFP/>, access date: 5 July 2023) was used to evaluate the allergenicity of the peptide and the output indicates the peptide is allergen and non-allergen [26–28]. Furthermore, ToxinPred (<https://webs.iitd.edu.in/raghava/toxinpred/design.php>, access date: 5 July 2023) was used to estimate the toxicity of the peptide [29]. The SVM (Swiss-Prot) + Motif-based method was set for the prediction method. Negative values indicate non-toxic classified results, while positive values indicate the possible toxicity of the analyzed peptide [30].

2.7. Assessment of Physicochemical Properties

The ExPASy ProtParam (<https://web.expasy.org/protparam/>, access date: 10 July 2023) tool was used to calculate the physicochemical properties of the conserved peptide,

which include molecular weight, theoretical isoelectric point (PI), instability index, amino acid composition, chemical formula, atomic composition, estimated half-life, aliphatic index, and grand average of hydropathicity (GRAVY) [31]. A negative GRAVY value indicates that the peptide is non-polar and hydrophilic. Also, a value below 40 for the instability index means the peptide is stable. Moreover, a higher aliphatic index shows higher thermostability. The solubility of the peptide was evaluated using the Pep-Calc protein calculator based on the PI, the number of charged residues, and the peptide length (<https://pepcalc.com/protein-calculator.php>, access date: 10 July 2023) [32]. The physico-chemical properties were used to estimate the stability, thermostability, and solubility of the selected conserved peptide.

3. Results

3.1. Selection of Conserved Peptide

The HA2 proteome of six highly pathogenic influenza virus strains among human populations was extracted from the NCBI database. The subtypes, NCBI accession number, and amino acid sequences of the HA2 domains are listed in Table 1. The multiple sequence alignment of the HA2 amino acid sequences showed two regions with the highest conservancy based on the conserved amino acid distribution (Figure 2, green and yellow boxes). The first 23 amino acids of the HA2 sequence (Figure 2, green box), which is known as fusion peptide, is embedded in the interspace between monomers of the HA trimer in the native conformation of the virus [33]. The second region with 20 amino acids in length (Figure 2, yellow box), HA_{288–107}: DVWTYNAELL VLMENERTLD, is not located at the virus membrane-associated region (Figure 3A). As the embedded region does not meet the peptide selection requirement for the current study, thus, only the HA_{288–107} sequence was selected for further analysis. The 3D structure of the HA protein and the position of the HA_{288–107} sequence are shown in Figure 3B. The conservancy analysis of the HA_{288–107} peptide showed an average of 84.16% identity for the peptide in the homologous protein sets (Table 2) and the NCBI BLAST analysis confirmed that the conserved peptide was 100% originated from the influenza virus. Thus, the HA_{288–107} sequence was selected for characterization and analysis as a potential antigen candidate.

Table 1. HA2 amino acid sequences of the highly pathogenic influenza A virus with their accession number, geographical regions, and date of collection.

| No. | Subtype | NCBI Accession Number | FASTA Format of HA2 Amino Acid Sequence |
|-----|---------|-----------------------|---|
| 1 | H1N1 | YP009121767 | >YP_009121767.1 FGAIAGFIEGGWTGMVDGWYGYHHQNEQSGYAADLKSTQ NAIDEITNKVNSVIEKMNTQFTAVGKEFNHLEKRIENLNKKVDD GFLDIWTYNAELLVLENERLTDYHDSNVKNLYEKVRSQKNN AKEIGNGCFEFYHKCDNTCMESVKNGTYDYPKYSEEAKLNREE IDGVKLESTRIYQILAIYSTVASSLVVSLGAIISFWMCSNGSLQCRICI |
| 2 | H3N2 | ASV62273 | >ASV62273.1 FGAIAGFIENGWEGMVDGWYGFRHQNSEGTGQAADLKSTQAAIN QITGKLNRIKKTNEKFHQIEKEFSEVEGRIQDLEKYVEDTKVDLWSY NAELLVALENQHTIDLTDSEMSKLFERTRRQLRENAEDMNGGCFKIY HKCDNACIGSIRNGTYDHDHYRDEALNNRFQIKGVQLKSGYKDWIL WISFAISCFLLCVLLGFIMWACQKGNIRCNICI |
| 3 | H5N1 | ACI06178 | >ACI06178.1 FGAIAGFIEGGWQGMVDGWYGYHHSNEQSGYAADKESTQKAID GVTNKVNSIIDKMNTQFEAVGREFNLERRIENLNKKMEDGFLDV WTYNAELLVLMENERTLDFHDSNVKNLYDKVRLQLRDNKELGN GCFEFYHKCDNECMESVRNGTYDYPQYSEEARLKREEISGVKLESIG IYQILSIYSTVASSLALAIMVAGLSLWMCSNGSLQCK |

Table 1. Cont.

| No. | Subtype | NCBI Accession Number | FASTA Format of HA2 Amino Acid Sequence |
|-----|---------|-----------------------|---|
| 4 | H7N9 | AGI60292 | >AGI60292.1 FGAIAGFIENGWEGLIDGWYGFRRHQNAQGEFTAADYKSTQ SAIDQITGKLNRLIEKTNQQFELIDNEFNEVEKQIGNVINWTRDSI TEVWSYNAELLVAMENQHTIDLADSEMDKLYERVKRQLRENAEE DGTGCFEIFHKCDDDCMASIRNNTYDHSKYREEAMQNRIQIDPV KLSSGYKDVILWFSFGASCFILLAIVMGLVFICVKNNGNMRTICI |
| 5 | H9N2 | CAB95857 | >CAB95857.1 FGAIAGFIEGGWVPLVAGWYGFQHSNDQGVGMAARDRSTQKAI DKITSKVNIVDKMNKQYEIIDHEFSEVETRLNMINNKIDDQIQDV WAYNAELLVLENQKTLDEHDANVNNLYNKVKRALGSNAMEEDG KGCFFELYHKCDDQCMETIRNGTYNRRKYREESRLERQKIEGVKL ESEGAYKILTIYSTVASSLVLAMGFAAFLFWAMSNGSCRCNICI |
| 6 | H2N2 | AAY28987 | >AAY28987.1 FGAIAGFIEGGWQGMVDGWYGYHHSNDQGSYAADKESTQK AFDGITNKVNSVIEKMNTQFEAVGKEFSNLERRLENLNKKMEDG FLDVWITYNAELLVLMENERTLDFHDSNVKNLYDKVRMQLRDNV KELGNGCFEFYHKCDDCECMNSVKNGTYDYPKYEEESKLNLRNEIK GVKLSSMGVYQILAIYATVAGSLSLAIMMAGISFWMCSNGSLQCRICI |

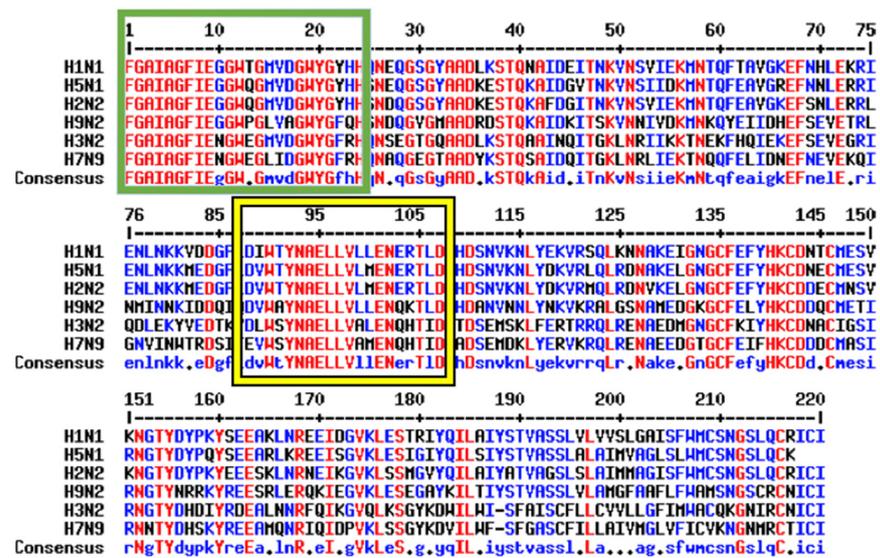


Figure 2. Multiple sequence alignment of HA2 amino acid sequences from six different strains of influenza virus using Multalin online software. The red color demonstrates the full conservancy, the blue color indicates one amino acid alternation and the black color represents the alternation in more than one amino acid at the specific region for different strains. The green (fusion peptide) and yellow (HA2_{88–107}) boxes show the regions with the highest conservation distribution. The letter in the consensus line means conservative and the dot indicates non-conservative amino acid substitution.

Table 2. HA2_{88–107} peptide identity among the six highly pathogenic strains of the influenza virus.

| HA2 _{88–107} Sequence | Peptide Identity among the Strains | | | | | |
|--------------------------------|------------------------------------|------|------|------|------|------|
| | H1N1 | H3N2 | H5N1 | H7N9 | H9N2 | H2N2 |
| DVWITYNAELLVLMENERTLD | 90% | 65% | 100% | 70% | 80% | 100% |
| | Average: 84.16% | | | | | |

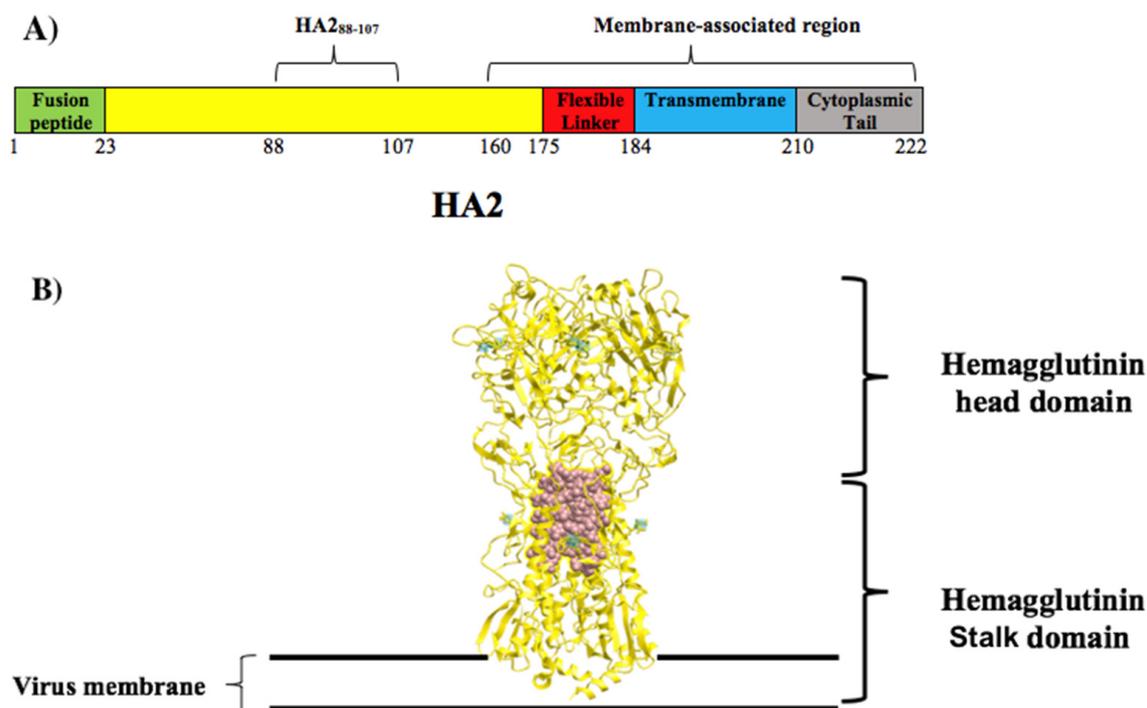


Figure 3. Location of the conserved HA2_{88–107} peptide on influenza HA protein. (A) Sequence diagram showing the polypeptide segment of the HA2 domain (222 amino acids) in its primary structure, indicating the HA2_{88–107} peptide is not associated with the membrane. (B) The position of HA2_{88–107} peptide (pink color) on the 3D full structure of HA protein (yellow color) using iCn3D online software, access date: 1 July 2023). The green color represents the carbohydrate components of the protein.

3.2. Linear B-Cell Epitopes of the HA2_{88–107} Peptide

Emini Surface Accessibility [18], Karplus and Schulz Flexibility [19], Chou and Fasman Beta-turn [20], Parker Hydrophilicity propensity [21], and Kolaskar and Tongaonkar Antigenicity analysis [22] methods from the IEDB server were used to identify linear B-cell epitopes on the HA2_{88–107} peptide (Figure 4). The amino acid sequence from 13 to 18 (*MENERT*) on the HA2_{88–107} peptide showed the highest surface accessibility with a score value of 2.745 (Figure 4A). The HA2_{88–107} peptide had a maximum Karplus and Schulz flexibility score value of 1.053 for the region between 13 to 19 amino acid residues (*MENERTL*) (Figure 4B). The Chou and Fasman secondary structure prediction tool identified an amino acid sequence from 1 to 7 (*DVWTYNA*) of the HA2_{88–107} peptide with the highest score value of 1.034 (Figure 4C). Moreover, Parker hydrophilicity analysis indicated that the most hydrophilic region of the HA2_{88–107} peptide is the amino acid sequence from 14 to 20 (*ENERTLD*) with a score value of 4.686 (Figure 4D). Kolaskar and Tongaonkar's analysis displayed that the sequence from 7 to 13 (*AELLVLM*) of the HA2_{88–107} peptide had the highest score value of 1.125 (Figure 4E). Generally, the higher the score value for the residues from the threshold might interpret that the amino acid sequence is having a higher probability to be part of an epitope (those residues are colored in yellow in Figure 4). In total, 20 linear B-cell epitopes were identified from which 5 of them were considered top-ranked epitopes based on the values and were tabulated in Table 3. The B-cell epitope identification results predicted that the HA2_{88–107} peptide is likely able to interact with B-cells and/or antibodies.

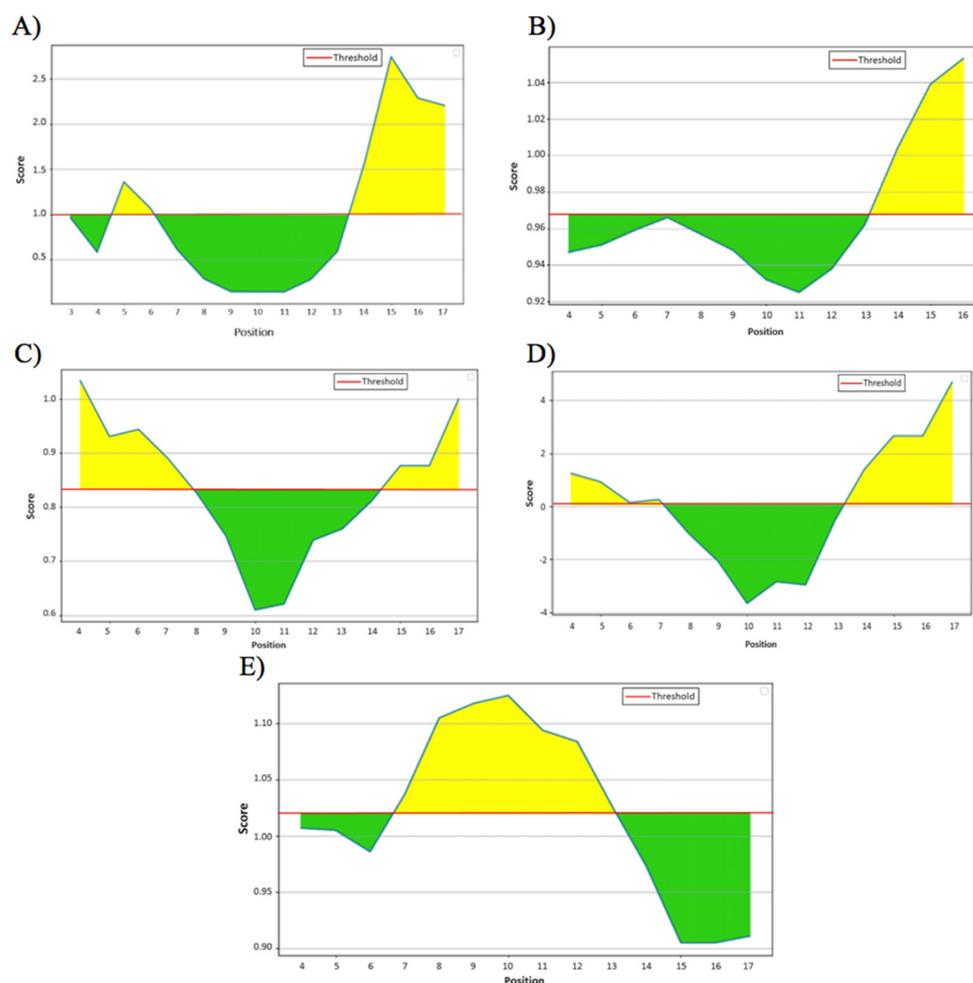


Figure 4. Linear B-cell epitope identification on HA₂₈₈₋₁₀₇ peptide. The most antigenic epitopes are shown in yellow color above the threshold value (red line) and green areas are not considered as an epitope. (A): Emini surface accessibility prediction with a threshold value of 1.000; (B): Karplus and Schulz flexibility prediction with a threshold value of 0.968; (C): Chou and Fasman secondary structure prediction with a threshold value of 0.834; (D): Parker hydrophilicity prediction with the threshold value of 0.064; (E): Kolaskar and Tongaonkar antigenicity prediction with the threshold value of 1.020.

Table 3. Top-ranked linear B-cell epitopes on HA₂₈₈₋₁₀₇ peptide.

| No. | Methods | Epitope Sequence | Start Position | End Position | Score Value | Threshold Value |
|-----|--------------------------------------|------------------|----------------|--------------|-------------|-----------------|
| 1 | Emini Surface Accessibility | MENERT | 13 | 18 | 2.745 | 1.000 |
| 2 | Karplus and Schulz Flexibility | MENERTL | 13 | 19 | 1.053 | 0.968 |
| 3 | Chou and Fasman Beta-turn | DVWPTYNA | 1 | 7 | 1.034 | 0.834 |
| 4 | Parker Hydrophilicity | ENERTLD | 14 | 20 | 4.686 | 0.064 |
| 5 | Kolaskar and Tongaonkar Antigenicity | AELLVLM | 7 | 13 | 1.125 | 1.020 |

3.3. T-Cell Epitopes of the HA₂₈₈₋₁₀₇ Peptide and Human Population Coverage

The T-cell epitope prediction tool from IEDB analyzed the MHC I and MHC II binding affinity of the HA₂₈₈₋₁₀₇ peptide with recommended methods. A total of 6 epitopes with a length of 9–11 mer each for MHC I and 6 epitopes with a length of 15 mer each for MHC II were obtained (Table 4). The MHC epitopes of the HA₂₈₈₋₁₀₇ were selected based on percentile ranks less or equal to 1% for MHC I and less and equal to 10% for MHC II [23]. These 12 epitopes showed binding affinity to a variety of MHC molecules and subsequently their

corresponding alleles, such as various Human Leukocyte Antigen (HLA) A and HLA-B for MHC I and HLA-DP, HLA-DQ, and HLA-DR for MHC II molecules. All the corresponding MHC alleles to the T-cell epitopes were used to calculate the global population coverage of MHC I and II alleles for the HA2_{88–107} peptide through the IEDB analysis server (Table 5). The MHC I alleles corresponding to T-cell epitopes cover more than 79.52% of the world population (Figure 5A). The average number of epitope hits/HLA combinations recognized by the world population was 3.12, and the minimum number recognized by 90% of the population (PC90) was 0.49 (Table 5). The MHC II alleles corresponding to the identified T-cell epitopes exhibited a population coverage rate of more than 99.88% (Figure 5B). The average number of epitope hits/HLA combinations recognized by the population was 13.79 and the PC90 was 7.81 (Table 5).

Table 4. T-cell epitopes on HA2_{88–107} peptide that bind to the human MHC class I and II alleles with high affinity.

| No. | Epitopes | MHC I Allele | Percentile Rank < 1 | Binding Energy (kcal/mol) | No. | Epitopes | MHC II Allele | Percentile Rank < 10 |
|---------------------------|------------|---|------------------------------------|---------------------------|-----|---------------------|---------------------------|----------------------|
| 1 | DVWTYNAEL | HLA-A*68:02 HLA-A*26:01 | 0.29 0.55 | −155.5 | 1 | DVWTYNA ELLVLMEN | HLA-DQA1*03:01/DQB1*03:02 | 3.7 |
| | | | | | | | HLA-DQA1*04:01/DQB1*04:02 | 4.6 |
| | | | | | | | HLA-DQA1*05:01/DQB1*02:01 | 4.8 |
| | | | | | | | HLA-DQA1*01:01/DQB1*05:01 | 6.3 |
| | | | | | | | HLA-DQA1*01:02/DQB1*06:02 | 9.4 |
| | | | | | | | HLA-DPA1*01:03/DPB1*04:01 | 3.8 |
| | | | | | | | HLA-DPA1*01:03/DPB1*02:01 | 4.4 |
| | | | | | | | HLA-DPA1*02:01/DPB1*01:01 | 5.2 |
| | | | | | | | HLA-DPA1*03:01/DPB1*04:02 | 7.8 |
| | | | | | | | HLA-DRB3*01:01 | 7.4 |
| HLA-DRB3*02:02 | 8.2 | | | | | | | |
| 2 | TYNAELLVL | HLA-A*24:02 HLA-A*23:01 | 0.08 0.15 | −103.2 | 2 | NAELLVLM ENERTLD | HLA-DRB1*04:05 | 1.9 |
| | | | | | | | HLA-DQA1*03:01/DQB1*03:02 | 2.2 |
| | | | | | | | HLA-DRB1*04:01 | 3.7 |
| | | | | | | | HLA-DRB1*01:01 | 4.9 |
| HLA-DRB1*15:01 | 6.7 | | | | | | | |
| 3 | TYNAELLVLM | HLA-A*24:02 HLA-A*23:01 | 0.43 0.48 | −115 | 3 | TYNAELLV LMENERT | HLA-DQA1*03:01/DQB1*03:02 | 1.5 |
| | | | | | | | HLA-DRB1*04:05 | 1.9 |
| | | | | | | | HLA-DQA1*04:01/DQB1*04:02 | 3.6 |
| | | | | | | | HLA-DRB1*04:01 | 3.7 |
| | | | | | | | HLA-DQA1*01:02/DQB1*06:02 | 6.5 |
| HLA-DRB1*15:01 | 8.5 | | | | | | | |
| 4 | VLMENERTL | HLA-A*02:01 HLA-A*02:03 HLA-A*02:06 HLA-A*32:01 HLA-B*08:01 | 0.1 0.15 0.3 0.46 0.51 | −160.8 | 4 | VWTYNAEL LVLME | HLA-DQA1*03:01/DQB1*03:02 | 1.3 |
| | | | | | | | HLA-DQA1*04:01/DQB1*04:02 | 3.1 |
| | | | | | | | HLA-DQA1*01:02/DQB1*06:02 | 4.2 |
| | | | | | | | HLA-DPA1*01:03/DPB1*02:01 | 4.4 |
| | | | | | | | HLA-DQA1*05:01/DQB1*02:01 | 4.4 |
| | | | | | | | HLA-DPA1*01:03/DPB1*04:01 | 4.6 |
| | | | | | | | HLA-DPA1*02:01/DPB1*01:01 | 5.8 |
| HLA-DPA1*03:01/DPB1*04:02 | 7.4 | | | | | | | |
| HLA-DRB3*01:01 | 9.6 | | | | | | | |
| 5 | VWTYNAELL | HLA-A*24:02 HLA-A*23:01 | 0.54 0.81 | −133.8 | 5 | WTYNAEL LVLME | HLA-DQA1*03:01/DQB1*03:02 | 1.5 |
| | | | | | | | HLA-DQA1*04:01/DQB1*04:02 | 3.5 |
| | | | | | | | HLA-DQA1*01:02/DQB1*06:02 | 4.6 |
| | | | | | | | HLA-DQA1*05:01/DQB1*02:01 | 4.7 |
| | | | | | | | HLA-DPA1*02:01/DPB1*01:01 | 7.9 |
| | | | | | | | HLA-DPA1*01:03/DPB1*04:01 | 8.6 |
| | | | | | | | HLA-DPA1*01:03/DPB1*02:01 | 10 |
| | | | | | | | HLA-DRB1*04:05 | 1.9 |
| HLA-DQA1*03:01/DQB1*03:02 | 2 | | | | | | | |
| 6 | WTYNAELLV | HLA-A*68:02 | 0.38 | −857.9 | 6 | YNAELLV LMENERTL | HLA-DRB1*04:01 | 3.7 |
| | | | | | | | HLA-DQA1*04:01/DQB1*04:02 | 5.4 |
| | | | | | | | HLA-DRB1*01:01 | 5.4 |
| | | | | | | | HLA-DRB1*15:01 | 7.9 |
| | | | | | | | | |

Table 5. Population coverage of the HA_{288–107} peptide for MHC I and II alleles.

| Population/Area | MHC Class I | | | MHC Class II | | |
|-----------------|-----------------------|--------------------------|-------------------|-----------------------|--------------------------|-------------------|
| | Coverage ^a | Average-Hit ^b | PC90 ^c | Coverage ^a | Average-Hit ^b | PC90 ^c |
| World | 79.52% | 3.12 | 0.49 | 99.88% | 13.79 | 7.81 |

^a projected population coverage. ^b average number of epitope hits/HLA combinations recognized by the population. ^c minimum number of epitope hits/HLA combinations recognized by 90% of the population.

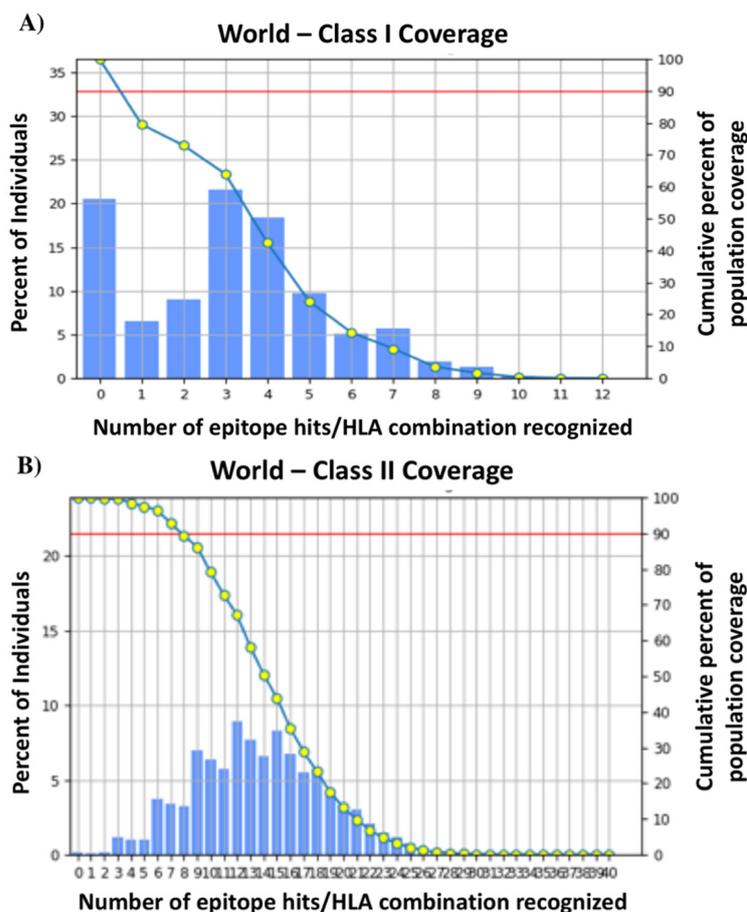


Figure 5. Population coverage of the HA_{288–107} peptide. The world populations were evaluated for the peptide using IEDB online population coverage analysis. (A) The graph shows a population coverage of 79.52% for MHC I epitopes. (B) The graph indicates population coverage of 99.88% for MHC II epitopes. The line (-o-) represents the cumulative percentage of population coverage of the epitopes; the bars represent the population coverage for each epitope; the red line represents PC90.

3.4. Cluster Analysis of the MHC-Restricted Alleles

The MHCcluster v2.0 online software was used to perform functional clustering of HLA molecules corresponding to the HA_{288–107} peptide based on correlations between predicted binding affinities [34]. The software produced a clustering of 9 HLA class I and 6 HLA class II, which were recognized to interact with the predicted epitopes and the output was produced through the conventional phylogenetic method based on sequence data available for different HLA alleles. Figure 6 illustrates the function-based clustering of the HLA alleles (heat map) with red zones demonstrating strong correlation and yellow zones indicating weaker interaction of the alleles.

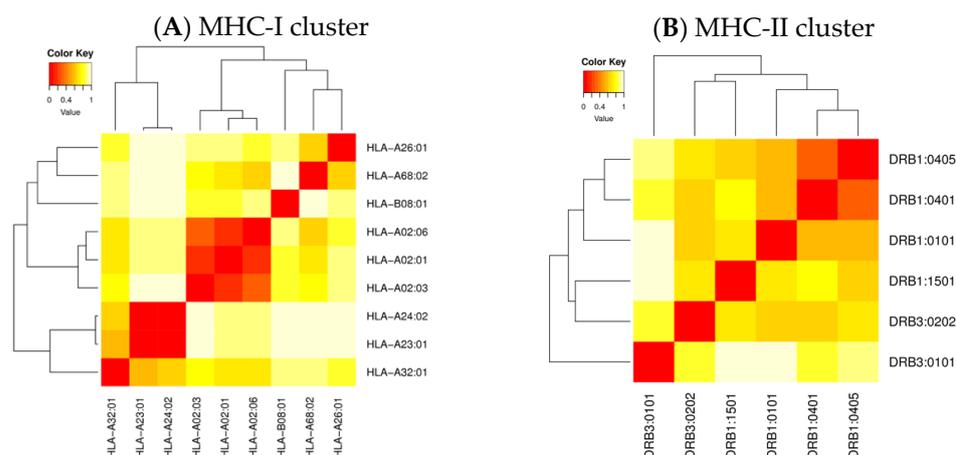


Figure 6. Cluster analysis of the HLA alleles. (A) Representing the cluster of MHC-I alleles and (B) Representing the cluster of MHC-II alleles. The red color indicates strong interaction and the yellow zone indicates weaker interaction with appropriate annotation.

3.5. Molecular Docking

The ClusPro software 2.0 was used for molecular docking analysis to investigate the binding pattern of the 6 MHC I predicted epitopes to the top-ranked corresponding HLA molecules (Table 4). Among the 6 epitopes, two epitopes, *VLMENERTL* and *WTYNAELLV*, interacted with the corresponding HLA molecules, HLA-A*02:01 and HLA-A*68:02, respectively, with high affinity at the binding groove of the molecules based on the lowest global energy and attractive Van Der Waals (VDW) in kcal/mol unit (Figure 7). The binding energy scores for MHC I molecules are shown in Table 4.

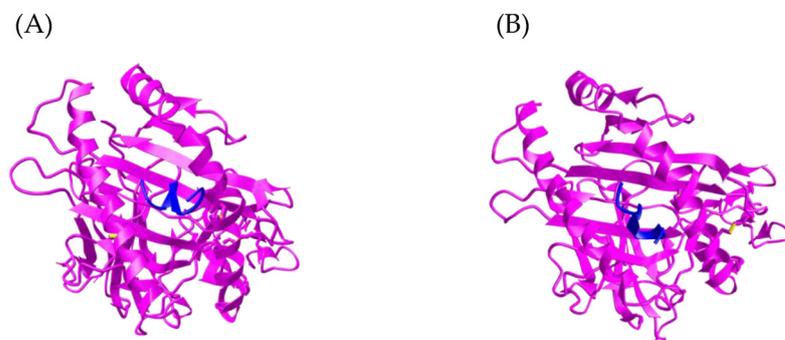


Figure 7. Visualization of ClusPro molecular docking of (A) *VLMENERTL* epitope interaction with HLA-A*02:01 molecule and (B) *WTYNAELLV* epitope interaction with HLA-A*68:02 molecule. Blue color indicates epitopes and magenta color indicates corresponding molecules (receptors).

3.6. Antigenicity and Safety of the HA_{288–107} Peptide

The total antigenicity prediction score for the HA_{288–107} peptide was 0.2370 via the VaxiJen 2.0 server which is below the threshold value of 0.4 which indicates the peptide has low antigenicity. In addition, the AllergenFP and ToxinPred outputs suggested the non-allergen and non-toxic nature of the peptide for humans, respectively.

3.7. Physicochemical Properties of the HA_{288–107} Peptide

In silico physicochemical analysis revealed that the molecular weight, PI value, and instability index of HA_{288–107} peptide were 2.42471 kDa, 3.83, and 1.99, respectively. The peptide is acidic and stable at pH 7 based on the PI and instability index values. Moreover, the result showed the peptide is 45% hydrophobic, 25% acidic, 5% basic, and 25% neutral. The peptide consists of 5 negatively charged residues and 1 positively charged residue. The chemical formula is C₁₀₇H₁₆₆N₂₆O₃₆S₁, and the total number of atoms is 336. The

half-life was estimated to be 1.1 h in mammalian reticulocytes, 3 min in yeast, and >10 h in *Escherichia coli*. The aliphatic index was estimated to be 112, indicating thermostability. The GRAVY was predicted to be -0.265 . A negative GRAVY value indicates that the peptide is non-polar and hydrophilic. The peptide was evaluated as a soluble protein in water based on the PI, the number of charged residues, and the peptide length. However, the only way to determine the exact solubility of a peptide is by laboratory experiments.

4. Discussion

The timely development of new vaccines is a scientific challenge to combat the ever-increasing global burden of viral diseases. In the case of the influenza virus, frequent antigenic changes in surface proteins, necessitate the reformulation of influenza vaccines every year [7]. So, the generation of a universal vaccine candidate that works against different pathogenic influenza virus strains could have significant public health and economic effects worldwide as it would surmount the current shortcomings of the annual vaccine and also could provide a cost-effective alternative to the annual influenza vaccine [35]. To achieve this goal, a conserved antigenic peptide needs to be identified. Thanks to advances in sequencing technology, we have abundant information about the genomics and proteomics of influenza virus. Therefore, it is possible to design peptide vaccines based on a conserved epitope using various immunoinformatic and cheminformatic tools [36,37].

In the past century, only H1N1, H2N2, and H3N2 influenza subtypes have caused pandemics, making these HA proteins the priority for vaccine development against future pandemics [38]. The second priority targets for vaccine development are defined by HA proteins from viruses that are known to have sporadically infected humans in the past such as H5N1, H7N9, and H9N2. Although second-group viruses are not easily transmitted to humans, they can cause very high mortality (30–60%) in the human population that is infected with these viruses [39,40]. In the present study, HA2 amino acid sequences of the mentioned six highly pathogenic influenza A viruses among the human population were retrieved from the NCBI database, and conserved regions were identified by multiple sequencing alignments. The HA2_{88–107} peptide was selected from two identified sequences as it was exposed on the surface of the virus compared to the other conserved regions based on 3D structure analysis. Also, it has been shown that the region between amino acids 35 and 107 of the HA2 sequence was accessible to the antibodies and immune cells in the native structure of the virus [41]. In general, antigens that are displayed on the surface of the virus are more likely accessible to the immune system at the first point of the infection [42]. Moreover, the HA2_{88–107} peptide was considered a highly conserved peptide as its average similarity (84.16%) with the homologous protein sets was more than 70%, showing that it may serve as a candidate antigen for the universal influenza vaccine design [43]. Although the H3N2 subtype showed 65% similarity to the HA2_{88–107} peptide (Table 2), the amino acid differences were conservative replacements, which means it is an exchange between two amino acids separated by a small physicochemical distance. It is necessary to check the similarity between the selected peptide and self-proteins while designing vaccine candidates to avoid any cross-reactivity of the induced immune responses [44]. The NCBI BLAST analysis confirmed that the peptide did not have any similarity to self-proteins or antigens from other pathogens and it 100% belongs to the influenza virus.

Epitope mapping has been performed to identify potential B and T-cell epitopes on the conserved HA2_{88–107} peptide to validate the antigen selection for vaccine design. An epitope refers to a specific site on an antigen to which a complementary antibody and/or immune cells may specifically bind. The B-cells, which are part of the adaptive immune system, act as receptors, bind to the antigen's epitopes, mediate humoral immunity, and have crucial roles in influenza viral infections and vaccinations [45,46]. In general, B-cell epitopes are categorized into linear (continuous) epitopes, which consist of a linear sequence of residues; and conformational (discontinuous) epitopes, which consist of residues that are brought together by folded protein structure and are not continuous in the primary protein sequence [47]. The prediction of linear epitopes has received major attention as

they can readily be used as antigens for vaccine and antibody production. In the case of conformational epitopes, 3D structure information and suitable scaffolds for epitope grafting are required. In the current study, only linear epitopes were predicted based on amino acid propensity scales depicting the physicochemical features of B-cell epitopes [48]. Twenty linear B-cell epitopes were identified that indicated the HA_{288–107} peptide is likely able to interact with B-cells. Among them, five epitopes are considered high-ranked epitopes based on surface accessibility, flexibility, secondary structure, hydrophilicity, and physicochemical properties.

The amino acid sequence from 13 to 18 (*MENERT*) on the HA_{288–107} peptide showed the highest surface accessibility with a score value of 2.745. The higher surface accessibility of the epitope means a higher chance of being recognized by B-cells or antibodies [18]. The HA_{288–107} peptide had a maximum flexibility score value of 1.053 for the region from 13 to 19 amino acid residues (*MENERTL*). When the epitope is flexible, several sets of conformations within the native antigen can cause the generation of different antibodies that lead to broader antibody-mediated responses [49]. The Chou and Fasman secondary structure analysis identified an amino acid sequence from 1 to 7 (*DVWTYNA*) with the highest score value of 1.034. The tool predicts the epitope based on the beta-turn structure scale, where the polypeptide chain folds back on itself by nearly 180 degrees and gives a protein its globularity rather than linearity structure and the beta turns in a protein are more surface accessible [20]. In general, the conformation of epitope allows recognition of the epitope by B-cells or antibodies, and in this case, it is thought to be an important aspect of vaccine design [50,51]. Furthermore, the Parker hydrophilicity analysis indicated that the most hydrophilic region of the HA_{288–107} peptide is the amino acid sequence from 14 to 20 (*ENERTLD*) with a score value of 4.686. As hydrophilic regions are predominantly located on the protein surface, subsequently, they can be considered as more antigenic regions [21,48]. The Kolaskar and Tongaonkar analysis displayed that the sequence from 7 to 13 (*AELLVLM*) of the HA_{288–107} peptide had the highest score value of 1.125. The method uses the physicochemical properties of amino acids and their frequencies of occurrence in experimentally known segmental epitopes to predict antigenicity [22].

On the other hand, cell-mediated immunity is another part of the adaptive immune response responsible for long-lasting immunity. It can restrict the spread of the infection directly and activate other immune cells to destroy pathogens, respectively [52,53]. Cytotoxic T-cells (CD8+) and helper T-cells (CD4+) play important roles in cell-mediated immunity. They scan other cells through their T-cell receptors for MHC:peptide complexes and recognize epitopes that are presented by MHC molecules [54]. When stimulated with a MHC molecule that presents antigenic peptides, T-cells can directly restrict the spread of infection in human cells and they have been able to provide cross-reactivity in the recognition of the different subtypes of influenza A virus when evoked by conserved regions of the influenza virus [55–57]. The epitope hits/MHC I molecule complex can trigger the CD8+ T cells and the epitope hits/MHC II molecule complex can be detected by CD4+ T cells, which is very important in the regulation of both CD8+ T and B-cells [58]. Thus, prediction of interaction between T-cell epitopes and MHC molecules is required in the design of vaccines against pathogens. Here, MHC binding prediction results demonstrated 12 T-cell epitopes on the HA_{288–107} peptide that can form MHC:epitope complexes to interact with CD8+ T and CD4+ T-cells and are considered potential epitope hits. These epitopes can interact with various MHC molecules in humans such as HLA-A and HLA-B for MHC class I and HLA-DP, HLA-DQ, and HLA-DR for MHC class II molecules. The epitopes were selected based on the percentile rank, which is a transformation that normalizes the binding affinity scores across different MHC molecules and enables MHC binding prediction and comparisons. A lower percentile rank value indicates higher affinity [23,59].

The MHC molecules are highly polymorphic and the expression frequency of different MHC molecules varies in different ethnicities and geographic areas around the world [60–62]. This polymorphism is basically in the gene region encoding the binding groove of MHC molecules and causes widely varying binding specificities [63]. Thus, any

selected epitope for epitope-based vaccine design should bind to several molecules of HLA supertype for maximum population coverage, especially in the case of universal vaccine development [43]. In general, population coverage is defined as the fraction of individuals in a population that responds to the predicted epitopes of an antigen based on the HLA allele frequencies of the population [64]. In the current study, all the corresponding alleles to the epitopes from T-cell epitope identification analysis were used to indicate the population coverage for the HA_{288–107} peptide. The results revealed a strong correlation between the corresponding MHC II alleles and the population coverage rate (99.88%). Thereby, the HA_{288–107} peptide can trigger CD4⁺ T cells and consequently, regulate CD8⁺ T and B-cell responses in 99.88% of the world population. Also, based on MHC I predicted alleles, the peptide can trigger CD8⁺ T cells directly in 79.52% of the world population.

Despite the high polymorphism in the human MHC genomic region, not all MHC molecules are equally different in terms of function. Structure-based clustering methods are effective in identifying superfamilies of MHC proteins with similar binding specificities [65]. As MHC super-families play an important role in drug development and vaccine design, MHC cluster analysis was also applied to specify the functional relationship of MHC variants based on correlations between predicted binding affinities [34]. The MHCcluster v2.0 server, which provides pictorial tree-based visualizations and a heat map of the functional association of MHC variants, generated a cluster of 15 HLA molecules identified as interacting with our predicted epitopes. The output was generated using a standard phylogenetic method based on available sequence data for different HLA alleles. The result of cluster analysis can indicate these alleles are in different cluster groups (more functionality in the population), which can strengthen the prediction from the previous analysis in this study. Furthermore, the molecular docking analysis in vaccine design can help researchers to model the interaction between epitopes and immune cell receptors at the atomic level and facilitate the prediction of the epitope binding possibilities to the immune receptors [66,67]. Hence, a docking study was performed to confirm the association between HLA molecules and our predicted epitopes. Two epitopes, *VLMENERTL* and *WTYNAELLV*, interacted with the corresponding alleles, HLA-A*02:01 and HLA-A*68:02, with high affinity at the binding groove of the HLA molecules. Consequently, the peptide is an ideal vaccine candidate for >1 billion people globally who express HLA-A*02:01 [68].

The VaxiJen 2.0 software classifies antigens based solely on the physicochemical properties of the protein and predicts the antigenicity of an antigen in total, irrespective of epitope mapping. The total antigenicity prediction score for the HA_{288–107} was below the threshold value of 0.4 which indicates the peptide has low antigenicity [28]. In general, one of the limitations of subunit vaccines is their low antigenicity as they mostly rely on epitopes with the absence of immunostimulatory components from the source pathogens. Different approaches such as the inclusion of nanoparticles and immunopotentiators into vaccine formulation have been applied to enhance the desired antigenicity and immunogenicity of influenza subunit vaccines [69]. For example, the Hepatitis B core Antigen was applied as an immunopotentiator agent to increase the HA_{288–107} peptide antigenicity [70]. Also, the result from VaxiJen analysis indicated that the total antigenicity score value was increased after the HA_{288–107} peptide was fused to the N terminal end of the HBcAg protein, which indicates the fusion protein can be antigenic.

The allergenicity assessment is a crucial step in the design of a peptide vaccine as it may induce allergic reactions by initiating immunoglobulin E production and enhancing response to histamine [71,72]. The result from AllergenFP software revealed that the HA_{288–107} peptide was not an allergen to humans so it should not be able to cause any allergic reaction within the body. The server is a binary classifier between allergens and non-allergens and it predicts allergenicity based on the amino acid principle properties such as size, hydrophobicity, helix, and beta-strand formations [73]. Moreover, high specificity and low cross-reactivity are generally considered to design effective, safe, and theoretically infallible therapeutic molecules and vaccines [30]. In particular, computational screening of non-toxic peptide approaches is needed to improve peptide selectivity with less cost and

time [74]. The ToxinPred tool was used to predict and evaluate the toxicity of the HA_{288–107} peptide in which the negative value indicated that the peptide was not toxic.

Various physicochemical properties of the peptide were analyzed. The instability index was computed below the threshold, which classified the peptide as a stable molecule and ensured that the construct possesses good characteristics to initialize an immunogenic reaction in the body. The value above the threshold indicates the instability of the protein, which can be related to the order of certain amino acids in the sequence [75]. On the other hand, the thermostability of the conserved peptide was demonstrated by the aliphatic index. The index is defined as the relative volume occupied by aliphatic side chains [alanine (Ala), valine (Val), isoleucine (Ile), and leucine (Leu)]. It may be regarded as a positive factor for the increase in thermostability of globular proteins. A high aliphatic index indicates that a protein is thermostable over a wide temperature range [76]. Moreover, the hydrophilic nature of the conserved peptide sequences was represented by the GRAVY value which is calculated as the sum of hydropathy values of all of the amino acids divided by the number of residues in the sequence. A negative GRAVY value indicates that the protein is non-polar and hydrophilic [77]. Furthermore, the PI value and the total number of charges indicated that theoretically, the peptide might dissolve in aqueous media at pH 7. In general, at the PI, proteins have a zero charge and tend to associate together, resulting in insolubility, but at the pH above the PI, the net charge is negative and solubility is likely predictable. On the other hand, the 45% hydrophobicity may cause insolubility or partial solubility of the peptide in an aqueous solution even if the sequence contains more than 25% charged residues [31,77]. So, further in vitro experiment is needed to discover the exact solubility of the conserved peptide.

5. Conclusions

In general, the influenza A virus remains a current and future threat to public health and all outbreaks demonstrate the need to develop a universal vaccine to protect against different strains. The computational framework is useful for evaluating the effectiveness of an antigen generated through existing bioinformatics tools. In the current study, the results indicated that the HA_{288–107} peptide can be a promising antigen for universal influenza vaccine design. However, we recommend further studies based on wet laboratory techniques to validate our predicted peptide experimentally for vaccine design and development.

Author Contributions: Conceptualization, A.K.-H.; methodology, A.K.-H.; formal analysis, A.K.-H.; writing—original draft preparation, A.K.-H.; writing—review and editing, A.K.-H., S.M., G.S., O.M. and Z.M.E.-B.; supervision, G.S.; project administration, A.K.-H.; funding acquisition, O.M., M.H.H., M.M.I. and Z.M.E.-B. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Deanship of Scientific Research at Northern Border University, Arar, KSA for funding this research work through the project number “NBU-FFR-2023-0128”.

Data Availability Statement: The data that support the findings of this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. World Health Organization. *World Health Organization Factsheet, Influenza (Seasonal)*; World Health Organization: Geneva, Switzerland, 2018.
2. Schild, G.C. The influenza virus: Antigenic composition and immune response. *Postgrad. Med. J.* **1979**, *55*, 87–97. [[CrossRef](#)] [[PubMed](#)]
3. Centers for Disease Control and Prevention (CDC); National Center for Immunization and Respiratory Diseases (NCIRD). 2019. Available online: <https://www.cdc.gov/flu/avianflu/influenza-a-virus-subtypes.htm> (accessed on 1 June 2023).
4. Peacock, T.P.; James, J.; Sealy, J.E.; Iqbal, M. A Global Perspective on H9N2 Avian Influenza Virus. *Viruses* **2019**, *11*, 620. [[CrossRef](#)]
5. Wilson, I.A.; Skehel, J.J.; Wiley, D.C. Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 Å resolution. *Nature* **1981**, *289*, 366–373. [[CrossRef](#)]

6. Xu, R.; Ekiert, D.C.; Krause, J.C.; Hai, R.; Crowe, J.E.; Wilson, I.A. Structural basis of preexisting immunity to the 2009 H1N1 pandemic influenza virus. *Science* **2010**, *328*, 357–360. Available online: <https://science.sciencemag.org/content/328/5976/357> (accessed on 1 June 2023). [[CrossRef](#)] [[PubMed](#)]
7. Subbarao, K.; Murphy, B.R.; Fauci, A.S. Development of Effective Vaccines against Pandemic Influenza. *Immunity* **2006**, *24*, 5–9. [[CrossRef](#)] [[PubMed](#)]
8. Skehel, J. An overview of influenza haemagglutinin and neuraminidase. *Biologicals* **2009**, *37*, 177–178. [[CrossRef](#)]
9. Hayati, M.; Biller, P.; Colijn, C. Predicting the short-term success of human influenza virus variants with machine learning. *bioRxiv* **2020**, *287*, 20200319. [[CrossRef](#)] [[PubMed](#)]
10. Blyth, C.C.; Macartney, K.K.; McRae, J.; Clark, J.E.; Marshall, H.S.; Buttery, J.; Francis, J.R.; Kotsimbos, T.; Kelly, P.M.; Cheng, A.C.; et al. Influenza Epidemiology, Vaccine Coverage and Vaccine Effectiveness in Children Admitted to Sentinel Australian Hospitals in 2017: Results from the PAEDS-FluCAN Collaboration. *Clin. Infect. Dis.* **2018**, *68*, eiy597. [[CrossRef](#)]
11. Wiley, S.K. The 2009 influenza pandemic: 10 years later. *Nursing* **2020**, *50*, 17–20. Available online: https://journals.lww.com/nursing/Fulltext/2020/01000/The_2009_influenza_pandemic_10_years_later.6.aspx (accessed on 1 June 2023). [[CrossRef](#)]
12. Erbeling, E.J.; Post, D.J.; Stemmy, E.J.; Roberts, P.C.; Augustine, A.D.; Ferguson, S.; Paules, C.I.; Graham, B.S.; Fauci, A.S. A Universal Influenza Vaccine: The Strategic Plan for the National Institute of Allergy and Infectious Diseases. *J. Infect. Dis.* **2018**, *218*, 347–354. [[CrossRef](#)]
13. Sah, P.; Alfaro-Murillo, J.A.; Fitzpatrick, M.C.; Neuzil, K.M.; Meyers, L.A.; Singer, B.H.; Galvani, A.P. Future epidemiological and economic impacts of universal influenza vaccines. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 20786–20792. [[CrossRef](#)]
14. Gori, A.; Longhi, R.; Peri, C.; Colombo, G. Peptides for immunological purposes: Design, strategies and applications. *Amino Acids* **2013**, *45*, 257–268. [[CrossRef](#)] [[PubMed](#)]
15. Moxon, R.; Reche, P.A.; Rappuoli, R. Editorial: Reverse Vaccinology. *Front. Immunol.* **2019**, *10*, 2776. [[CrossRef](#)]
16. Wu, N.C.; Wilson, I.A. A Perspective on the Structural and Functional Constraints for Immune Evasion: Insights from Influenza Virus. *J. Mol. Biol.* **2017**, *429*, 2694–2709. [[CrossRef](#)]
17. 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](#)]
18. 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](#)]
19. Karplus, P.A.; Schulz, G.E. Prediction of chain flexibility in proteins. *Sci. Nat.* **1985**, *72*, 212–213. [[CrossRef](#)]
20. Chou, P.Y.; Fasman, G.D. Prediction of the Secondary Structure of Proteins from their Amino Acid Sequence. *Adv. Enzymol. Relat. Areas Mol. Biol.* **1978**, *47*, 45–148. [[CrossRef](#)] [[PubMed](#)]
21. Parker, J.M.R.; Guo, D.; Hodges, R.S. New hydrophilicity scale derived from high-performance liquid chromatography peptide retention data: Correlation of predicted surface residues with antigenicity and x-ray-derived accessible sites. *Biochemistry* **1986**, *25*, 5425–5432. [[CrossRef](#)] [[PubMed](#)]
22. 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](#)]
23. Reynisson, B.; Alvarez, B.; Paul, S.; Peters, B.; Nielsen, M. NetMHCpan-4.1 and NetMHCIIpan-4.0: Improved predictions of MHC antigen presentation by concurrent motif deconvolution and integration of MS MHC eluted ligand data. *Nucleic Acids Res.* **2020**, *48*, W449–W454. [[CrossRef](#)]
24. Sharma, S.; Kumari, V.; Kumbhar, B.V.; Mukherjee, A.; Pandey, R.; Kondabagil, K. Immunoinformatics approach for a novel multi-epitope subunit vaccine design against various subtypes of Influenza A virus. *Immunobiology* **2021**, *226*, 152053. [[CrossRef](#)]
25. Li, L.; Dong, M.; Wang, X.G. The implication and significance of beta 2 microglobulin: A conservative multifunctional regulator. *Chin. Med. J.* **2016**, *129*, 448–455. [[CrossRef](#)]
26. Dimitrov, I.; Bangov, I.; Flower, D.R.; Doytchinova, I. AllerTOP v.2—A server for in silico prediction of allergens. *J. Mol. Model.* **2014**, *20*, 2278. [[CrossRef](#)] [[PubMed](#)]
27. Dimitrov, I.; Naneva, L.; Doytchinova, I.; Bangov, I. AllergenFP: Allergenicity prediction by descriptor fingerprints. *Bioinformatics* **2014**, *30*, 846–851. [[CrossRef](#)] [[PubMed](#)]
28. 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](#)]
29. Gupta, S.; Kapoor, P.; Chaudhary, K.; Gautam, A.; Kumar, R.; Raghava, G.P.S. In Silico Approach for Predicting Toxicity of Peptides and Proteins. *PLoS ONE* **2013**, *8*, 9. [[CrossRef](#)]
30. Harnkit, N.; Khongsonthi, T.; Masuwan, N.; Prasartkul, P.; Noikaew, T.; Chumnanpuen, P. Virtual Screening for SARS-CoV-2 Main Protease Inhibitory Peptides from the Putative Hydrolyzed Peptidome of Rice Bran. *Antibiotics* **2022**, *11*, 1318. [[CrossRef](#)] [[PubMed](#)]
31. Gasteiger, E.; Hoogland, C.; Gattiker, A.; Duvaud, S.; Wilkins, M.R.; Appel, R.D.; Bairoch, A. *Proteomics Protocols Handbook*; Humana Press: Totowa, NJ, USA, 2005.
32. Lear, S.; Cobb, S.L. Pep-Calculator: A set of web utilities for the calculation of peptide and peptoid properties and automatic mass spectral peak assignment. *J. Comput. Aided. Mol. Des.* **2016**, *30*, 271–277. [[CrossRef](#)]
33. Skehel, J.J.; Wiley, D.C. Receptor Binding and Membrane Fusion in Virus Entry: The Influenza Hemagglutinin. *Annu. Rev. Biochem.* **2000**, *69*, 531–569. [[CrossRef](#)]

34. Thomsen, M.; Lundegaard, C.; Buus, S.; Lund, O.; Nielsen, M. MHCcluster, a method for functional clustering of MHC molecules. *Immunogenetics* **2013**, *65*, 655–665. [CrossRef]
35. Kelsey, C.; Mhatre, V.M.; Ho, J.-A.L.; Martin, R.; Buchwald, S.L.; Ho, M.V.; Martin, K.C.; Craik, J.-A.L.; Manuscript, C.; Kantrowitz, A. NIH Public Access. *Bone* **2008**, *23*, 1–7.
36. Knutson, K.L.; Schiffman, K.; Disis, M.L. Immunization with a HER-2/neu helper peptide vaccine generates HER-2/neu CD8 T-cell immunity in cancer patients. *J. Clin. Investig.* **2001**, *107*, 477–484. [CrossRef] [PubMed]
37. López, J.A.; Weilenman, C.; Audran, R.; Roggero, M.A.; Bonelo, A.; Tiercy, J.; Spertini, F.; Corradin, G. A synthetic malaria vaccine elicits a potent CD8⁺ and CD4⁺ T lymphocyte immune response in humans. Implications for vaccination strategies. *Eur. J. Immunol.* **2001**, *31*, 1989–1998. [CrossRef] [PubMed]
38. Nabel, G.J.; Fauci, A.S. Induction of unnatural immunity: Prospects for a broadly protective universal influenza vaccine. *Nat. Med.* **2010**, *16*, 1389–1391. [CrossRef] [PubMed]
39. Webster, R.G.; Govorkova, E.A. Continuing challenges in influenza. *Ann. N. Y. Acad. Sci.* **2014**, *1323*, 115–139. [CrossRef]
40. Widdowson, M.-A.; Bresee, J.S.; Jernigan, D.B. The Global Threat of Animal Influenza Viruses of Zoonotic Concern: Then and Now. *J. Infect. Dis.* **2017**, *216* (Suppl. 4), S493–S498. Available online: <https://pubmed.ncbi.nlm.nih.gov/28934463> (accessed on 1 June 2023). [CrossRef]
41. Stepanova, L.A.; Sergeeva, M.V.; Shuklina, M.A.; Shaldzhyan, A.A.; Potapchuk, M.V.; Korotkov, A.V.; Tsybalova, L.M. A fusion protein based on the second subunit of hemagglutinin of influenza A/H2N2 viruses provides cross immunity. *Acta Nat.* **2016**, *8*, 116–126. [CrossRef]
42. Maginnis, M.S. Virus–Receptor Interactions: The Key to Cellular Invasion. *J. Mol. Biol.* **2018**, *430*, 2590–2611. [CrossRef]
43. Eickhoff, C.S.; Terry, F.E.; Peng, L.; Meza, K.A.; Sakala, I.G.; Van Aartsen, D.; Moise, L.; Martin, W.D.; Schriewer, J.; Buller, R.M.; et al. Highly conserved influenza T cell epitopes induce broadly protective immunity. *Vaccine* **2019**, *37*, 5371–5381. [CrossRef]
44. Sette, A.; Rappuoli, R. Reverse Vaccinology: Developing Vaccines in the Era of Genomics. *Immunity* **2010**, *33*, 530–541. [CrossRef] [PubMed]
45. Chen, X.; Liu, S.; Goraya, M.U.; Maarouf, M.; Huang, S.; Chen, J.-L. Host Immune Response to Influenza A Virus Infection. *Front. Immunol.* **2018**, *9*, 320. [CrossRef] [PubMed]
46. Paul, W.E. *Fundamental Immunology* Wolters Kluwer Health; *Fundamental Immunology*. 2012. Available online: <https://books.google.com.my/books?id=x0gpapT-Y2kC> (accessed on 1 June 2023).
47. Jespersen, M.C.; Mahajan, S.; Peters, B.; Nielsen, M.; Marcatili, P. Antibody Specific B-Cell Epitope Predictions: Leveraging Information From Antibody-Antigen Protein Complexes. *Front. Immunol.* **2019**, *10*, 298. Available online: <https://www.frontiersin.org/article/10.3389/fimmu.2019.00298> (accessed on 1 June 2023). [CrossRef]
48. Sanchez-Trincado, J.L.; Gomez-Perosanz, M.; Reche, P.A. Fundamentals and Methods for T- and B-Cell Epitope Prediction. *J. Immunol. Res.* **2017**, *2017*, 2680160. [CrossRef]
49. Gori, A.; Peri, C.; Quilici, G.; Nithichanon, A.; Gaudesi, D.; Longhi, R.; Gourlay, L.; Bolognesi, M.; Lertmemongkolchai, G.; Musco, G.; et al. Flexible vs Rigid Epitope Conformations for Diagnostic- and Vaccine-Oriented Applications: Novel Insights from the *Burkholderia pseudomallei* BPSL2765 Pal3 Epitope. *ACS Infect. Dis.* **2016**, *2*, 221–230. [CrossRef] [PubMed]
50. Pellequer, J.-L.; Westhof, E.; Van Regenmortel, M.H. Correlation between the location of antigenic sites and the prediction of turns in proteins. *Immunol. Lett.* **1993**, *36*, 83–99. [CrossRef] [PubMed]
51. Malonis, R.J.; Lai, J.R.; Vergnolle, O. Peptide-Based Vaccines: Current Progress and Future Challenges. *Chem Rev.* **2020**, *120*, 3210–3229. [CrossRef]
52. Igietseme, J.U.; Eko, F.O.; He, Q.; Black, C.M. Antibody regulation of T-cell immunity: Implications for vaccine strategies against intracellular pathogens. *Expert Rev. Vaccines* **2004**, *3*, 23–34. [CrossRef]
53. Bacchetta, R.; Gregori, S.; Roncarolo, M.-G. CD4⁺ regulatory T cells: Mechanisms of induction and effector function. *Autoimmun. Rev.* **2005**, *4*, 491–496. [CrossRef]
54. Patronov, A.; Doytchinova, I. T-cell epitope vaccine design by immunoinformatics. *Open Biol.* **2013**, *3*, 120139. Available online: <http://www.ncbi.nlm.nih.gov/pubmed/23303307> <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3603454> (accessed on 1 June 2023). [CrossRef]
55. Knapp, B.; van der Merwe, P.A.; Dushek, O.; Deane, C.M. MHC binding affects the dynamics of different T-cell receptors in different ways. *PLoS Comput. Biol.* **2019**, *15*, e1007338. [CrossRef]
56. McMichael, A.J.; Gotch, F.M.; Noble, G.R.; Beare, P.A.S. Cytotoxic T-Cell Immunity to Influenza. *N. Engl. J. Med.* **1983**, *309*, 13–17. [CrossRef]
57. Hoft, D.F.; Babusis, E.; Worku, S.; Spencer, C.T.; Lottenbach, K.; Truscott, S.M.; Abate, G.; Sakala, I.G.; Edwards, K.M.; Creech, C.B.; et al. Live and Inactivated Influenza Vaccines Induce Similar Humoral Responses, but Only Live Vaccines Induce Diverse T-Cell Responses in Young Children. *J. Infect. Dis.* **2011**, *204*, 845–853. [CrossRef]
58. Gustiananda, M. Immunoinformatics analysis of H5N1 proteome for designing an epitope-derived vaccine and predicting the prevalence of pre-existing cellular-mediated immunity toward bird flu virus in Indonesian population. *Immunome Res.* **2011**, *7*, 1.
59. Moutaftsi, M.; Peters, B.; Pasquetto, V.; Tschärke, D.C.; Sidney, J.; Bui, H.-H.; Grey, H.; Sette, A. A consensus epitope prediction approach identifies the breadth of murine TCD8⁺-cell responses to vaccinia virus. *Nat. Biotechnol.* **2006**, *24*, 817–819. [CrossRef] [PubMed]

60. Maleki, A.; Russo, G.; Palumbo, G.A.P.; Pappalardo, F. In silico design of recombinant multi-epitope vaccine against influenza A virus. *BMC Bioinform.* **2022**, *22*, 617. [[CrossRef](#)] [[PubMed](#)]
61. Lokki, M.; Paakkanen, R. The complexity and diversity of major histocompatibility complex challenge disease association studies. *HLA* **2019**, *93*, 3–15. [[CrossRef](#)]
62. Reche, P.A.; Reinherz, E.L. Sequence Variability Analysis of Human Class I and Class II MHC Molecules: Functional and Structural Correlates of Amino Acid Polymorphisms. *J. Mol. Biol.* **2003**, *331*, 623–641. [[CrossRef](#)]
63. Bui, H.-H.; Sidney, J.; Dinh, K.; Southwood, S.; Newman, M.J.; Sette, A. Predicting population coverage of T-cell epitope-based diagnostics and vaccines. *BMC Bioinform.* **2006**, *7*, 477–484. [[CrossRef](#)]
64. González-Galarza, F.F.; Takeshita, L.Y.; Santos, E.J.; Kempson, F.; Maia, M.H.T.; Da Silva, A.L.S.; Silva, A.L.T.E.; Ghattaoraya, G.S.; Alfievic, A.; Jones, A.R.; et al. Allele frequency net 2015 update: New features for HLA epitopes, KIR and disease and HLA adverse drug reaction associations. *Nucleic Acids Res.* **2015**, *43*, D784–D788. [[CrossRef](#)]
65. Larsen, M.E.; Kloverpris, H.; Stryhn, A.; Koofhethile, C.K.; Sims, S.; Ndung'u, T.; Goulder, P.; Buus, S.; Nielsen, M. HLArestrictor—A tool for patient-specific predictions of HLA restriction elements and optimal epitopes within peptides. *Immunogenetics* **2010**, *63*, 43–55. [[CrossRef](#)]
66. McConkey, B.J.; Sobolev, V.; Edelman, M. The performance of current methods in ligand–protein docking. *Curr. Sci.* **2002**, *83*, 845–856. Available online: <http://www.jstor.org/stable/24107087> (accessed on 1 June 2023).
67. Zahroh, H.; Ma'rup, A.; Tambunan, U.S.F.; Parikesit, A.A. Immunoinformatics approach in designing epitopebased vaccine against meningitis-inducing bacteria (*Streptococcus pneumoniae*, *Neisseria meningitidis*, and *Haemophilus influenzae* type b). *Drug Target Insights* **2016**, *10*, 19–29. [[CrossRef](#)] [[PubMed](#)]
68. Valkenburg, S.A.; Venturi, V.; Dang, T.H.Y.; Bird, N.L.; Doherty, P.C.; Turner, S.J.; Davenport, M.P.; Kedzierska, K. Early Priming Minimizes the Age-Related Immune Compromise of CD8⁺ T Cell Diversity and Function. *PLoS Pathog.* **2012**, *8*, e1002544. Available online: <https://pubmed.ncbi.nlm.nih.gov/22383879> (accessed on 1 June 2023). [[CrossRef](#)]
69. Khalaj-Hedayati, A.; Chua, C.L.L.; Smooker, P.; Lee, K.W. Nanoparticles in influenza subunit vaccine development: Immunogenicity enhancement. *Infl. Other Respir. Viruses* **2019**, *14*, 92–101. [[CrossRef](#)] [[PubMed](#)]
70. Khalaj-Hedayati, A.; Chua, C.L.L.; Smooker, P.; Lee, K.W. Universal influenza vaccine technologies and recombinant virosome production. In *Methods in Recombinant Protein Production*, 1st ed.; Elsevier Ltd.: Amsterdam, The Netherlands, 2022; pp. 45–89. [[CrossRef](#)]
71. Odelram, H.; Granström, M.; Hedenskog, S.; Duchén, K.; Björkstén, B. Immunoglobulin E and G responses to pertussis toxin after booster immunization in relation to atopy, local reactions and aluminium content of the vaccines. *Pediatr. Allergy Immunol.* **1994**, *5*, 118–123. [[CrossRef](#)] [[PubMed](#)]
72. Sen, D.K.; Arora, S.; Gupta, S.; Sanyal, R.K. Studies of adrenergic mechanisms in relation to histamine sensitivity in children immunized with Bordetella pertussis vaccine. *J. Allergy Clin. Immunol.* **1974**, *54*, 25–31. [[CrossRef](#)]
73. McKeever, T.M.; Lewis, S.A.; Smith, C.; Hubbard, R. Vaccination and Allergic Disease: A Birth Cohort Study. *Am. J. Public Health* **2004**, *94*, 985–989. [[CrossRef](#)]
74. Robles-Loaiza, A.A.; Pinos-Tamayo, E.A.; Mendes, B.; Ortega-Pila, J.A.; Proaño-Bolaños, C.; Plisson, F.; Teixeira, C.; Gomes, P.; Almeida, J.R. Traditional and Computational Screening of Non-Toxic Peptides and Approaches to Improving Selectivity. *Pharmaceuticals* **2022**, *15*, 323. [[CrossRef](#)]
75. Guruprasad, K.; Reddy, B.; Pandit, M.W. Correlation between stability of a protein and its dipeptide composition: A novel approach for predicting in vivo stability of a protein from its primary sequence. *Protein Eng. Des. Sel.* **1990**, *4*, 155–161. [[CrossRef](#)]
76. Gasteiger, E.; Hoogland, C.; Gattiker, A.; Duvaud, S.; Wilkins, M.R.; Appel, R.D.; Bairoch, A. Protein Identification and Analysis Tools on the ExPASy Server. In *The Proteomics Protocols Handbook*; Walker, J.M., Ed.; Humana Press: Totowa, NJ, USA, 2005; pp. 571–607. [[CrossRef](#)]
77. Zayas, J.F. (Ed.) *Solubility of Proteins BT—Functionality of Proteins in Food*; Springer: Berlin/Heidelberg, Germany, 1997; pp. 6–75. [[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.