Vaccine Research
ISSN: 2423-4923
eISSN: 2383-2819

Computational Design and Prokaryotic Expression of a Hemagglutinin-Based Influenza Vaccine: Preliminarily in vitro and *in vivo* Study

Soheila Hoseini [1, 2], Parastoo Ehsani [2], Fatemeh Fotuhi [3], Behrokh Farahmand [3*], Houshang Alizadeh [1*]

[1]Department of Agricultural Biotechnology, Agricultural College, Tehran University, Karaj Campus, Karaj 31587-77871, Iran. [2]Department of Molecular Biology, Pasteur Institute of Iran, 69 Pasteur Ave., Tehran 1316943551, Iran. [3]Department of Influenza and Other Respiratory Viruses, Pasteur Institute of Iran, 69 Pasteur Ave., Tehran 1316943551, Iran.

ARTICLE INFO

Research Article

VacRes, 2023
Vol. 10, No.2, 28 – 38
Received: September 26, 2024
Accepted: January 05, 2025
Pasteur Institute of Iran

***Corresponding Authors:**

Houshang Alizadeh;
Department of Agricultural Biotechnology, Agricultural College, Tehran University, Karaj Campus, Karaj 31587-77871, Iran
Email: halizade@ut.ac.ir
Tel/Fax: +98-2632224022

Behrokh Farahmand;
Department of Influenza and Other Respiratory Viruses, Pasteur Institute of Iran, 69 Pasteur Ave., Tehran 1316943551, Iran
Email: b_farahmand@pasteur.ac.ir
Tel/Fax: +98-2164112181

KEYWORDS: Influenza, *In silico*, Multi-epitope, Recombinant vaccine

ABSTRACT

Introduction: Influenza A (H1N1) virus is one of the main causes of seasonal and pandemic influenza in humans. Considering the inherent antigenic changes of the virus, new infectious strains of the virus emerge frequently. In addition, the current methods of vaccine production are mainly based on fully inactivated viruses, which are time-consuming and expensive approaches. In this regard, the development of new protective vaccines will be a reasonable approach. This study aimed to design and express a subunit vaccine based on the hemagglutinin protein of the pH1N1 strain, enriched with protective epitopes of four highly pathogenic virus strains, namely, H1N1, H3N2, H5N1, and H7N9. **Methods:** Using bioinformatics tools, highly antigenic and conserved epitopes were selected from H1N1, H3N2, H5N1, H7N9 strains and placed at the appropriate positions on HA protein framework of pH1N1 influenza virus. The *in silico* analyses including immunogenicity, conservancy and population coverage were used to examine the construct, named as Hemag98. The genes sequence was synthesized and transferred to an *Escherichia coli* host for expression. The expression of Hemag98 protein was then confirmed by SDS-PAGE and Western blotting techniques. The immunogenicity of the purified recombinant proteins was preliminary examined in a mouse model. **Results:** Based on the predicted results, Hemag98 was shown to be a stable and hydrophilic protein with similar tertiary structure to native HA1 of H1N1 virus. Based on the Z-score of the predicted models, the Hemag98 model was in range of X-ray structured proteins. The expressed protein had the expected 45 KDa molecular weight and the initial *in vivo* experiment showed that the recombinant Hemag98 protein stimulates specific immune responses. **Conclusion:** The overall results indicated that the Hemag98 can be considered as a promising platform for the development of a preventive influenza vaccine capable of protecting against diverse strains of influenza virus.

INTRODUCTION

Influenza is a highly contagious respiratory disease and a significant global health concern. According to the Centers for Disease Control and Prevention (CDC) in the USA, between 2010 and 2020, the influenza caused millions of illnesses, hundreds of thousands of hospitalizations, and tens of thousands of deaths, annually[1]. Influenza viruses are RNA viruses that belong to the Orthomixoviridae family which possess immunogenic surface proteins, such as hemagglutinin (HA). There are four types of influenza viruses, namely A, B, C, and D. While all three types of influenza viruses can cause epidemics, only type A may lead to pandemics, and it is assigned to a higher proportion of human infectivity. Influenza viruses undergo minor (antigenic drift) and major (antigenic shift) antigenic changes due to their error-prone RNA-dependent RNA polymerase and segmented genome. These changes enable the virus to evade adaptive immune response in various avian and mammalian host species, including humans[2]. The primary prophylactic measure to reduce the burden of influenza is annual vaccination.

Currently, influenza vaccines are composed of three virus strains selected annually. There are four types of influenza vaccines available: (a) whole virus vaccines that contain either inactivated or live-attenuated viruses; (b) split virus vaccines consisting of virus fragments; (c) subunit vaccines or purified antigens, containing the surface proteins HA and Neuraminidase (NA); and (d) virosomal vaccines comprising synthetic virus-like particles with embedded HA and NA virus surface proteins. All of these vaccine types are designed to target specific strains, and

[DOI: 10.52547/vacres.10.2.28]

their effectiveness hinges on their ability to stimulate the production of antibodies that match neutralizing epitopes on the HA protein of the anticipated strains, likely to infect during the upcoming influenza season[3].

According to the CDC data, commercially available strain-specific vaccines have shown relatively poor efficacy, ranging between 10% - 60% among the general population, even when there is a good match between a vaccine and the circulating strains. Furthermore, current vaccines do not safeguard against newly-emerging influenza subtypes, which have resulted in multiple pandemics. Therefore, there is an urgent need to explore alternative approaches to traditional influenza vaccines with low efficacy[4]. Universal influenza vaccines represent a viable solution as they can offer broad and long-lasting protection against a variety of virus strains without the need for annual updates. Consequently, this subject has become a significant area of research in the field of influenza vaccines. New strategies for developing universal vaccines are focused on selecting conserved epitopes or proteins of influenza that are shared among seasonal and pre-pandemic strains, such as matrix protein, nucleoprotein, and HA stalk domain[5].

Traditional vaccines have been developed by injecting attenuated or killed disease-causing agents. However, there has been a recent revolution in genome and protein sequencing, which has led to the discovery of a novel approach, known as Reverse Vaccinology and Vaccinomics. This rapid *in silico* method highlights that vaccine design is possible from sequence information without the need to cultivate the pathogens. The strategy aims to combine human immunology and immunogenetics with bioinformatics to discover and design new vaccines, such as universal influenza vaccines[6].

Advancements in reverse genetics techniques have reached a stage where it is now feasible to generate influenza viruses that carry foreign epitopes within their viral proteins. This enables the introduction of foreign epitopes to the immune system using hybrid influenza virus proteins. To accomplish this, it is crucial to identify the specific regions in viral proteins that can accept the insertion or substitution of different epitopes. The HA glycoprotein of the influenza virus plays a vital role in triggering neutralizing antibodies and is a prime target for the vaccine development. After infection, a majority of antibodies produced against the influenza virus are directed towards particular regions within the HA protein. These regions, referred to as hypervariable antigenic sites A, B, C, D, and E, reside in the head region of the HA protein. Due to their significance, these antigenic domains within the HA protein offer favorable locations for incorporating foreign epitopes[7].

Previous researches have focused on creating vaccines using the HA protein in different ways. These include expressing the protein alone from a specific strain, making a fusion HA protein with adjuvant proteins, or multi-epitopes, composed of conserved epitopes connected together using linkers. In this study, we used a different approach in which the HA protein was utilized as a scaffold for developing a multi-epitope vaccine without the need for linkers. Previous studies conducted by Li et al. (1991) and Zhu et al. (2005) have shown that the HA protein, has highly variable regions within its head domain which, can effectively be replaced with epitopes or even larger proteins without compromising its structure and function[7, 8]. Therefore, we designed here a modified influenza HA antigen, called Hemag98 through computational approaches and expressed it in an *E. coli* expression system. This antigen included conserved epitopes from the HA and M2e proteins of common circulating (H1N1, H3N2) as well as potentially dangerous (H5N1, H7N9) strains of influenza viruses, replaced into the head domain of the HA protein.

MATERIALS AND METHODS

Ethics Statement

Animal experiments were approved and performed based on the guidelines introduced by the Ethics Committee of the Pasteur Institute of Iran (IR.PII.REC.1399.059). BALB/c mice were housed in a well-lighted (12 h light/dark cycle), air-conditioned room (26 ± 1 °C) with 50 ± 10% of humidity. They had free access to standard diet and water.

Primary Sequence Collection

First, we retrieved the complete amino acid sequences of influenza HA from 2015-2020 in FASTA format from NCBI. We performed sequence alignment using Clustal Omega (http://www.ebi.ac.uk/Tools/msa/Clustalo) to identify conserved regions. Then, we used CD-Hit and the cons tab of EMBOSS tools (https://www.bioinformatics.nl/emboss-explorer/) to obtain the final consensus sequence.

B-cell Epitopes Prediction

B-cell epitopes are specific regions on an antigen that can bind to the receptors, found on the surface of B lymphocytes. Identifying and characterizing these epitopes is crucial for designing effective vaccines. To predict linear B-cell epitopes, various tools from IEDB were utilized. These tools utilize sequence characteristics of the antigen, such as hydrophilicity (Parker Hydrophilicity Prediction), flexibility (Karplus and Schulz flexibility scale), accessibility (Emini surface accessibility scale), polarity, and antigenicity (Kolaskar and Tongaonkar antigenicity scale). Additionally, BepiPred-2.0 was employed to predict linear B-cell epitopes.

T-cell Epitope Prediction

To identify MHC class I and II epitopes for consensus sequences, we employed the IEDB-recommended method from their website (http://tools.iedb.org/mhci/mhcii). The Consensus method was used in combination with Artificial Neural Network (ANN), Stabilized Matrix Method (SMM), and Combining Library (CombLib) predictors, if corresponding predictors were available for the molecule [2]. If not, NetMHCpan and NetMHCIIpan were used instead [3]. The server allowed multiple allele/length pairs to be submitted at once, using peptide lengths of 9 and 15 for MHCI and MHCII, respectively. We also included frequently occurring alleles that are present in at least 1% of the human population in the allele box. To further specify the prediction for the Iranian population, we added the Iranian allele set to our analysis.

Epitope Conservancy and Human Homology

In an epitope-based vaccine framework, utilizing conserved epitopes can provide broader protection. To assess the conservancy of selected epitopes across multiple strains, we used the IEDB conservancy tool (http://tools.iedb.org/conservancy/) with a sequence identity threshold above 70%. As a final step, we filtered the selected epitopes by comparing their homology to the human proteome using the BLAST tool (https://blast.ncbi.nlm.nih.gov/Blast.cgi) available on NCBI. Selected epitopes incorporated into the head domain of the HA protein of pH1N12009 (accession number: AHZ40793). Computational chemistry and bioinformatics techniques were

[DOI: 10.52547/vacres.10.2.28]

employed to assess the quality, stability, and potential immunogenicity of the designed vaccine construct (*i.e.*, Hemag98).

Vaccine Candidate Analyses

SOPMA server was used to predict Secondary structure of Hemag98 (https://npsaprabi.ibcp.fr/NPSA/npsa_sopma.html). This alignment-based server predicts a three-state description of the secondary structure (alpha-helix, beta-sheet, and coil) [9]. Tertiary structure of Hemag98 was predicted using ROBETTA server (https://robetta.bakerlab.org) [10]. The ROBETTA server, which is a CASP-certified protein structure prediction server [10], which model proteins based on *ab initio* and comparative modeling (CM) methods. In the CM method, the protein sequence is divided into domains and models and each domain is built using homology sequences to proteins of known structure. Unaligned regions are treated as loops[10].

To assess the quality of protein models, ProSA web online tool was used[11]. The ProSA-web Z-score plot displays the Z-scores of all experimentally determined protein chains in the current Protein Data Bank. Light blue and dark blue represent the Z-scores of X-ray and NMR-derived structures, respectively[11]. In order to compute various physical and chemical parameters of Hemag98 protein, including molecular weight (Mw), isoelectric pH (pI), aliphatic index, GRAVY (grand average hydropathy) values, hydropathicity, instability index, and estimated half-life, the EXPASY ProtParam tool (http://expasy.org/cgi-bin/protparam) was used[12]. As a final step, ANTIGENpro and SOLpro from the Scratch Protein Predictor (http://scratch.proteomics.ics.uci.edu) were used to predict the whole protein antigenicity and solubility overexpression in *E. coli*[13]. The allergenicity of the protein was carried out by the AlgPred server (http://www.imtech.res.in/raghava/algpred/submission.html) [14].

Codon Optimization and Cloning of the Vaccine Construct

The protein construct was codon-optimized for expression in *E. coli* by online servers (ExpOptimizer and OPTIMIZER)[15]. Finally, restriction sites were inserted at both ends of the sequence and synthesized by Biomatics Company (Canada). The gene was inserted into *Bam*HI and *Xho*I sites of a pET28a vector with double digestion, and transformation of the *E. coli* cells was done by the heat shock protocol[16]. Selection of the transformed bacteria was done by culturing the bacteria on a selective LB agar, containing kanamycin (50 µg/ml). To prove the insertion of the gene of interest in the vector, double digestion with previously mentioned restriction enzymes, as well as PCR, and DNA sequencing were performed on the extracted plasmids. *E. coli* Top10 and BL21 (DE3) strains as well as female BALB/c mice aged 6–8 weeks old were supplied by the Pasteur Institute of Iran, Tehran, Iran. Isopropyl b-D-1-Thiogalactopyranoside (IPTG), Tetramethylbenzidine (TMB), Horseradish Peroxidase (HRP) conjugated anti-mouse Immunoglobulin G (IgG), HRP-conjugated goat anti-rabbit IgG, and also 96-well plates were obtained from Sigma-Aldrich Company (USA). Alum adjuvant was prepared from Alhydrogel 2% (Brenntag Biosector, Denmark; CAS Number: 21645-51-2).

Recombinant Protein Expression and Electrophoresis

Confirmed recombinant expression vector pET28a-mHA1 was transformed into *E. coli* BL21 (Novagen, USA). Recombinant colonies were sub-cultured in LB broth at 37°C until reaching OD_{600} of 0.4-0.7. The cultures were then induced with various IPTG (Sigma-Aldrich Company, USA) concentrations (0.5, 1, 2, 4 mM) and temperature (20, 28, 37°C, O/N), evaluated by SDS-PAGE. Scale-up production of the recombinant protein was performed in LB broth medium inoculated with a single colony containing a recombinant vector. The bacterial suspensions were then centrifuged (4°C, 10,000 rpm, 10 min), and the pellets were stored at - 20°C. The Lysis-Equilibration-Wash (LEW) buffer, urea (8 M), and sonication (20 pulses, 80–90% pulse rate, 20 s on and off intervals) were used for the protein extraction.

Purification and Confirmation of Recombinant Protein

The recombinant protein had two 6X-histidine tags at N/C terminals and was purified through the application of a Ni-NTA affinity column (Qiagen, Germany). Briefly, a pellet from 250 ml of induction medium (1mM IPTG) was resuspended in 10 mL of lysis buffer, and the cells were sonicated on ice 20 times (with 20-second pulses/relaxation intervals). The lysates were loaded onto the Ni-NTA column and incubated for 1 h at 4°C with gentle shaking. The column was washed with wash buffer containing different imidazole concentrations (30, 50, 100, 200 mM) to remove nonspecifically bound proteins. In the final step, the desired recombinant protein was eluted from the column by applying elution buffer (250 mM imidazole). The purified protein was dialyzed against phosphate-buffered saline (PBS) (pH 8.0), with the buffer being changed every 1h to remove urea and imidazole. Protein concentration was calculated using the BCA kit (DNAbiotech). The purity of the recombinant protein was shown in 12% SDS-PAGE and approved with a Western blotting assay. Expression of recombinant Hemag98 in both reduced and non-reduced form was further assayed with Western blotting and ELISA, respectively. For Western blotting, the proteins were first separated by SDS-PAGE and then transferred to nitrocellulose paper using a semi-dry transfer apparatus (Bio-Rad). The transferred proteins were incubated for 1 h in PBS buffer containing a 1:6000 dilution of HRP-conjugated anti-His antibody (Qiagen, Germany), also home-made anti-HA1(1:100) which previously produced and purified by Khosravi *et al* in influenza department of Pasture institute of Iran [17], as first antibody and HRP conjugated goat anti-rabbit (1:50000) as secondary antibody used. The bands were visualized by exposure to diaminobenzidine (DAB) substrate. For ELISA, the striped plate was coated overnight with 100 ng/ well of Hemag98 along with control proteins at 4°C. Home-made anti-HA1 were diluted as 1:1000, HRP-conjugated goat anti-rabbit (1:50000) was used as secondary antibody and the optical density was measured at 450 nm wavelength.

Immunization Procedure of Mice

This experimental research was carried out at the Pasture Institute of Iran, in 2021.

Immunogenicity of vaccine protein (Hemag98) was evaluated in BALB/c mice. The animals were randomly divided into 2 case groups (5 mice/each group), and the animals individually received Hemag98 supplemented with Alum/Alhydrogel (3mg/ml; 1:1). The mice that received PBS were considered vehicle control groups. The compounds were administered subcutaneously in a total volume of 100 µl containing 15 µg of the Hemag98 protein for three times at 15-day interval. Before immunization and two weeks after the last injection, the mice

[DOI: 10.52547/vacres.10.2.28]

were bled through the orbital sinus and the sera were kept at − 20 °C for further use.

Measurement of Specific Anti- Hemag98 Antibodies

Blood samples were collected from 5 mice of each group, two weeks after each immunization, and the blood sera were used to measure specific IgG antibodies using the ELISA technique. Briefly, 96-well ELISA plates were coated overnight at 4 °C using 100 µl of 20ng/well of the Hemag98 antigen. Sera of the mice were diluted in PBS (1:1000- 1:512000). The concentration of coated antigen was obtained by a checkerboard titration assay. Phosphate buffer saline with a pH of 7.4 (10 mM) containing 0.05% of Tween 20 (PBS-T) was used for washing steps and PBS-T containing 5% of bovine serum albumin was used as blocking buffer to prevent non-specific binding of antigens and antibodies to a microtiter plate. To evaluate specific IgG, HRP-conjugated goat anti-mouse IgG (1:5000, Sigma) was used as a secondary antibody and optical density was measured at 450 nm of wavelength.

Statistics

All statistical analyses were carried out by GraphPad Prism version
9 (GraphPad Software Inc., USA). One-way ANOVA was applied to compare the differences between the study groups. The $P < 0.05$ was considered significant.

RESULTS

From Epitope Prediction to Modeling and Validation

The protein sequence of HA from the Influenza type A virus database was retrieved from NCBI. A sequence alignment was performed to identify the highly conserved regions of the sequence among all strains from 2015-2020. The selected final epitopes were integrated in the head domain of pH1N1 2009 while preserving the original receptor binding site (RBS) sequence of pH1N1. According to the SOPMA result, the protein model consisted of 17.71% helix, 20.5% extended strand, 7.55% beta turn, and 54% random coil. Based on the results obtained from the ROBETTA server, the models created using the CM method were found to be better than those generated by the *ab initio* method, based on the error estimate point of view. As previously mentioned, both CM prediction methods with and without custom templates were employed. The server created 5 models, all of which were very similar. Based on the error estimate graph provided for all predicted models, the best model was created using the database alignment of the server (Fig. 1) which had a lower error estimate, compared to those made using a custom template (6WCR).

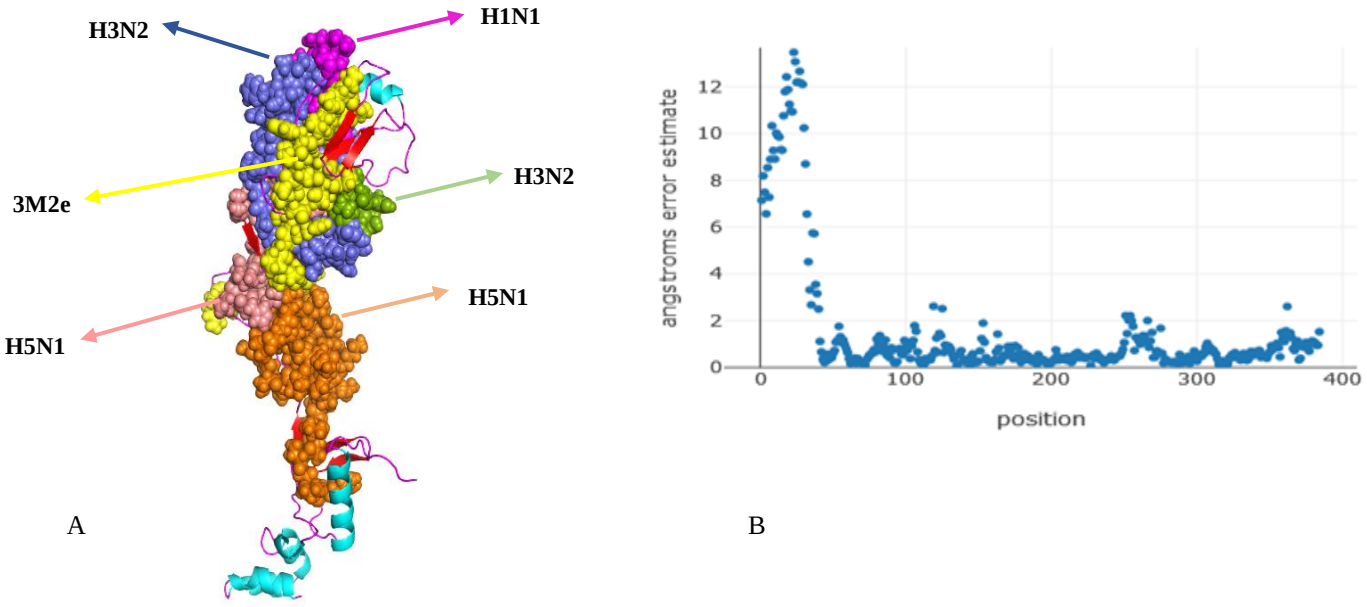

Fig. 1. (A) 3D model predicted by ROBETTA server, the incorporated epitopes were indicated by arrows. (B) Plot of error estimate for predicted 3D model which indicated low errors (<2 Aº).

ProSA-web analysis and Ramachandran plots (https://zlab.umassmed.edu/bu/rama/index.pl) of the Hemag98 model are shown in Fig. 2B. As presented, more than 94% of the residues are located in the highly preferred zone, and 5.2% are in the preferred region. No outliers (0.0%) were observed from the same plot, confirming the correct design and folding of the protein. The given Z-score for our model by the server was -5.68 (Fig. 2A), which falls within the range of proper conformations based on X-ray crystallography. Furthermore, the quality of the molecular structure was evaluated by Protein Structure.
Validation Software (PSVS) and the results showed that the protein model is acceptable (Table1).

[DOI: 10.52547/vacres.10.2.28]

Table 1. Structural indices of Hemag98 protein using PSVS.

Program	Verify3D	ProsaII (-ve)	Procheck (phi-psi)	Procheck (all)	MolProbity Clashscore
Row score	0.22	0.19	-0.53	-0.21	8.65
Z-score	-3.85	-1.9	-1.77	-1.24	0.04

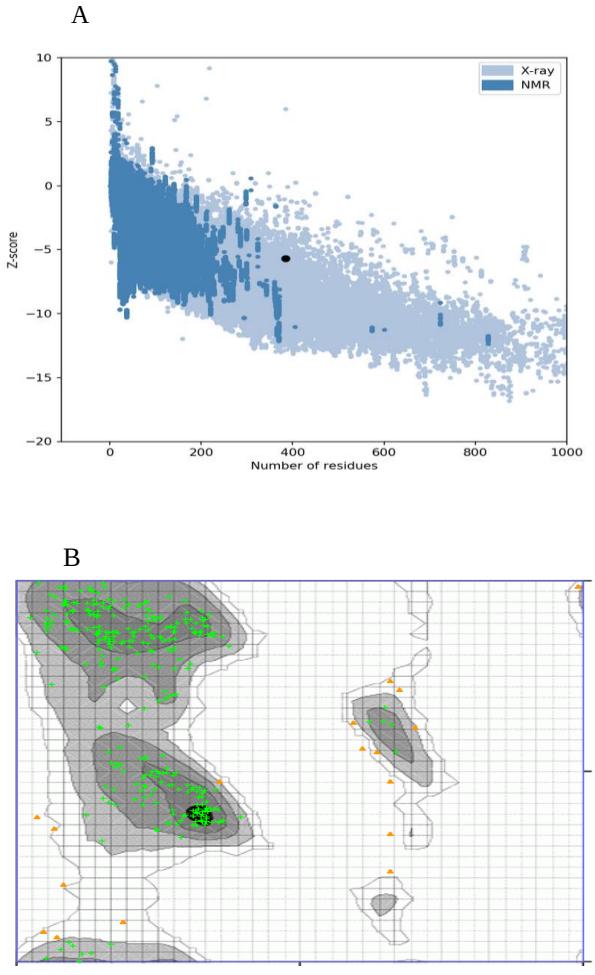

Fig. 2. (A) Z-score plot created by Prosa Web server, model Hemag98 (black dot) is in the acceptable region for X-ray based models. (B) The Ramachandran plot of model Hemag98, green dots are residues of Hemag98 model protein which mostly (94%) located in highly preferred region (dark grey).

Physiochemical Analysis of the Vaccine Construct

ProtParam is a tool that computes various physicochemical properties deduced from a protein sequence[12]. The molecular weight of the protein was approximately 42 kDa, with a theoretical pI of 7.6. The half-life of our protein was estimated to be more than 10 hours, based on the N-terminal amino acid sequence. This represents the prediction of the time it takes for half of the amount of protein in a cell to disappear after its synthesis in the cell. The instability and aliphatic index estimated by ProtParam indicate the stability and relative volume occupied by aliphatic side chains, respectively. Based on the results obtained from this server, the protein was found to be stable due to its relatively small instability index and higher aliphatic index. The Grand Average of Hydropathicity (GRAVY) was determined to be -0.411, indicating that Hemag98 belongs to the hydrophilic group. The antigenicity of the whole protein was assessed using VaxiJen server and ANTIGENpro. Hemag98 scored 0.45 and 0.938462 in VaxiJen and ANTIGENpro, respectively. The results from both servers were high enough and above the threshold (≥0.4 in VaxiJen) to conclude that the protein is antigenic. SOLpro was used to predict the propensity of the protein to be soluble upon overexpression in *E. coli*. The predicted value was 0.959729, indicating that it is likely to be soluble. Allergen analysis of the protein also confirmed that the designed protein is not classified as an allergenic protein.

Epitope and Population Coverage Prediction of the Final Vaccine Structure

To ensure the exposure of the epitopes, online server for linear B cell epitopes and MHCI/II bindings sites prediction was used. The results confirm the presence and exposure of the epitopic sites. The epitopes with high scores are shown in Table 2 and Table 3.

Table 2. Linear B-cell epitope predicted by ElliPro

NO.	Start	End	Peptide	Number of residues	Score
1	64	109	VNLLEDKHNGKLCKKHLLSRINHFEKIQIIPKSSWSSHEASLGVR	45	0.88
2	119	142	SDNGTCSKAYSNCYPYDVPDYAS	23	0.79
3	152	176	TLEFINEGFNWTGVTQNGGSSACGA	25	0.72
4	188	204	GKNGLYPKLSQTYINDK	17	0.62
5	279	310	LTEVETPTRGPGPGSLLTEVETPIRDTPCHDC	32	0.54
6	319	330	GAINSSMPFHNI	12	0.72
7	342	362	VKQNTLKLATGLRNVPSIQSR	21	0.612

[DOI: 10.52547/vacres.10.2.28]

Table 3. Prediction of BALB/c mice MHC I/II binding sites.

Number	Allele	Peptide	length	start	Percentile rank	Affinity score
	MHCI					
1	H-2-Db	TGLRNVPSI	9	308	0.01	0.8799
2	H-2-Kb	VTVTHSVNL	9	8	0.04	0.5921
3	H-2-Dd	RGPGPGSLL	9	236	0.01	0.53633
4	H-2-Dd	IGPGPGSLL	9	220	0.01	0.72251
5	H-2-Kb	GAKSFYKNL	9	124	0.11	0.52039
6	H-2-Db	YVKQNTLKL	9	290	0.11	0.4218
7	H-2-Db	MPFHNIHPL	9	274	0.15	0.3492
8	H-2-Db	SFYKNLIWL	9	127	0.19	0.2919
9	H-2-Db	KAYSNCYPY	9	76	0.32	0.278
10	H-2-Kb	IWLTGKNGL	9	133	1.5	0.172
11	H-2-Kb	TGLRNVPSI	9	300	0.29	0.456
12	H-2-Dd	VASSGTLEF	9	96	0.31	0.3335
13	H-2-Db	SLLTEVETY	9	210	1.3	0.1123
14	H-2-Dd	IQIIPKSSW	9	40	0.78	0.483
	MHCII					
1	H2-IAd	DTVLEKNVTVTHSVN	15	1	0.79	0.709
2	H2-IAd	TVLEKNVTVTHSVNL	15	2	1.1	0.6602
3	H2-IAd	VPDYASLRSLVASSG	15	86	2	0.5634

Population coverage result from IEDB server, also was used to assay the coverage of the proposed vaccine in Iranian population based on the allele frequency. Fig. 3 represents approximately 98% and 60.5% coverage based on the MHCI and MHCII allele, respectively.

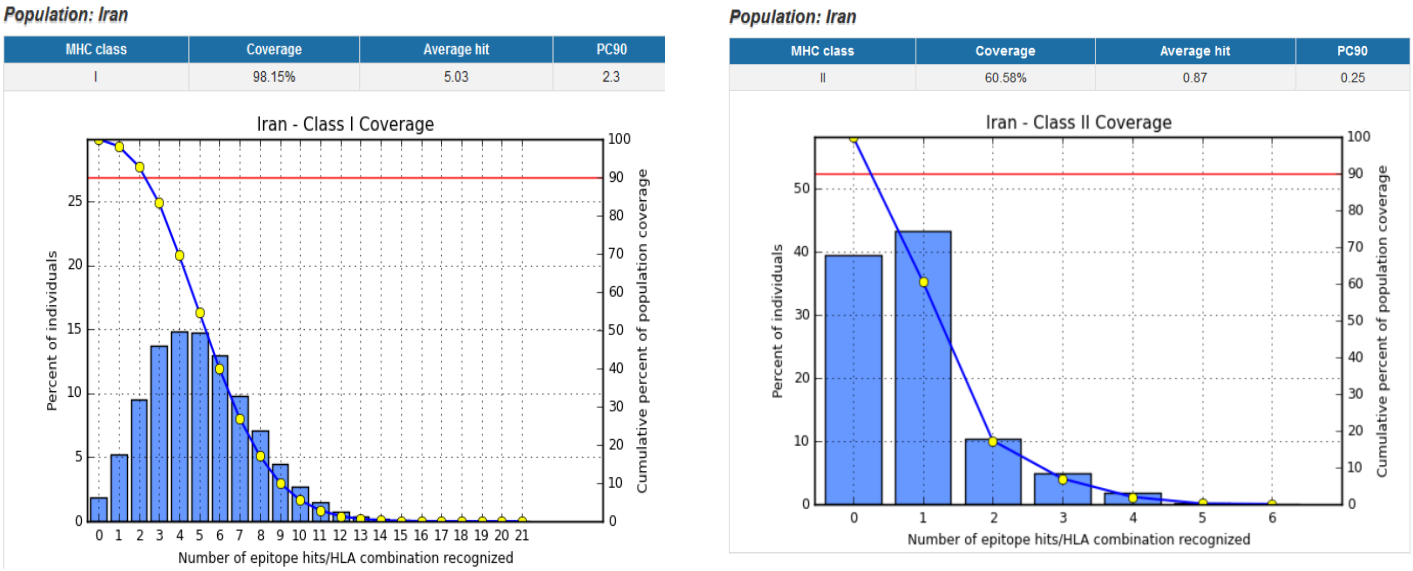

Fig 3. Population coverage prediction of the designed construct by IEDB server in Iranian population.

Codon Optimization and Cloning of Hemag98 Protein

The codon-optimized sequence was checked for possible amino acid changes using two servers. No changes were found in the protein sequence. The calculated Codon Adaptation Index (CAI) increased to 1.0 (OPTIMIZER) and 0.84 (ExpOptimizer) after optimization based on the *E. coli* codon table usage. Additionally, the GC content of the optimized sequence was found to be 53%.

Cloning and Transformation of Vaccine Construct

The gene fragment encoding Hemag98 was successfully inserted into the pET28a vector (Fig. 4), and the presence of Hemag98 in the vector was confirmed through various methods such as restriction digestion, PCR, and DNA sequencing (Fig. 5). The recombinant vectors were then extracted from *E. coli* TOP10 cells and transformed into the expression host, *E. coli* BL21. Successful transformation was confirmed via colony-touch PCR.

[DOI: 10.52547/vacres.10.2.28]

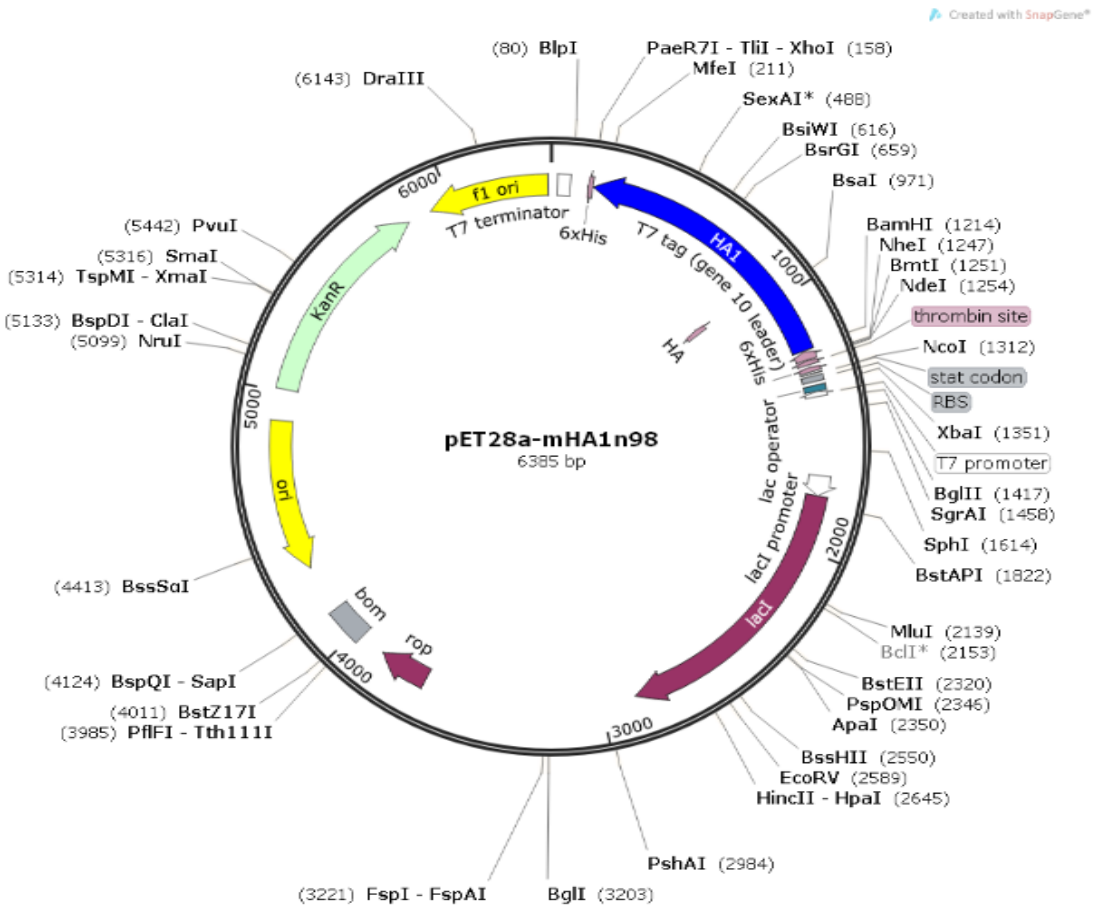

Fig. 4. Schematic presentation of cloned Hemag98 in pET28a vector.

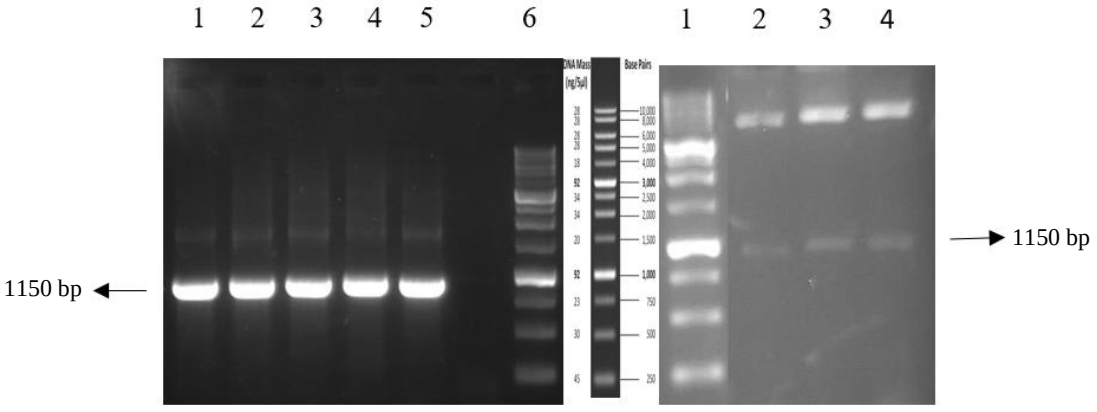

Fig. 5. Cloning confirmation of pET28- Hemag98. **(A)** Colony-touch PCR result, lane(1-5) and lane (6) DNA ladder; the band with desired size amplified in all 5 colons. **(B)** Double digestion of pET28-Hemag98 plasmids [figure (A) clonies (2-4)] with *Bam*HI and *Xho*I and lane(1) is DNA lader.

[DOI: 10.52547/vacres.10.2.28]

Vaccine Protein Expression Optimization

The expressed recombinant protein was visualized via 12% SDS-PAGE (Fig. 6). Based on results from protein expression optimization, 1mM IPTG along with 37 ºC and 4 hours for induction, culture temperature and induction time, respectively, were best condition. Following protein purification, a band of approximately 42 kDa, consistent with bioinformatic predictions, was observed on SDS-PAGE (Fig. 6). The identity of the recombinant protein was subsequently confirmed using western blotting analysis after incubation with both anti-His and anti-HA1 antibodies. The immunoblotting results of purified Hemag98, using home-made anti-HA1 indicated that the structure of recombinant protein have not been impaired following prokaryotic system expression and denaturing purification method.

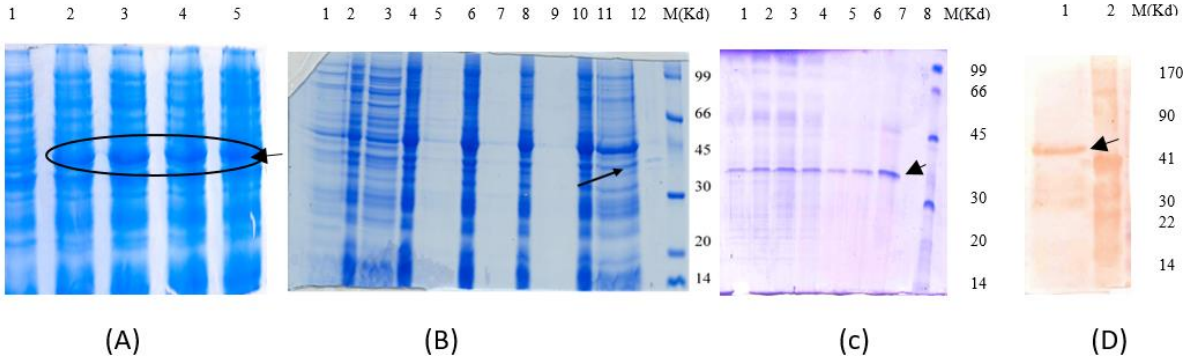

Fig 6. (A) SDS-PAGE result of Hemag98 protein expression in *E.coli*; lane (1) before induction, lane (2-5) after induction using IPTG (0.5mM). **(B)** SDS-PAGE result of Hemag98 protein extraction, lanes (1, 3, 5, 7, 9 and 11) are supernatant and lanes (2, 4, 6, 8, 10 & 12) are pellets' wells**. (C)** SDS-PAGE result of Hemag98 protein purification using imidazole (30-300mM), lane(1); imidazole 30mM; lanes (2-3) imidazole 50mM; and lane (4); imidazole 100mM, lanes (5-6) imidazole 200mM, lane (7); imidazole 250mM and lane(8); imidazole 300. **(D)** Confirmation of produced recombinant Hemag98 with anti- HA1 antibody, lane (1) purified protein, lane (2); bacterial extract, and lane (3) protein marker (the arrows indicate the position of Hemag98 protein).

Measurement of Specific Antibodies against Recombinant Vaccine

The potency of the recombinant protein in production of specific antibodies was evaluated, using a mouse model. The protein was administered in combination with Alum adjuvant. The mice were bled and specific antibody responses against Hemag98 recombinant proteins were evaluated using an ELISA, two weeks after the first, second, and third immunizations. Blood samples obtained prior to immunization were considered as the negative control. IgG Antibody titer increased significantly after the third injection, compared to the first and the second immunizations (Fig. 7 A). The results of the last immunization indicated that the titer of anti-Hemag98 in the test groups, compared to the control serum, was detectable at dilutions up to 1:128,000. This demonstrates the recombinant protein's strong ability to induce high levels of antibody production (Fig. 7 B). Based on the *P*-value (0.0008), there are significant differences between each vaccinated group and the control group (PBS-injected; Fig. 7 C).

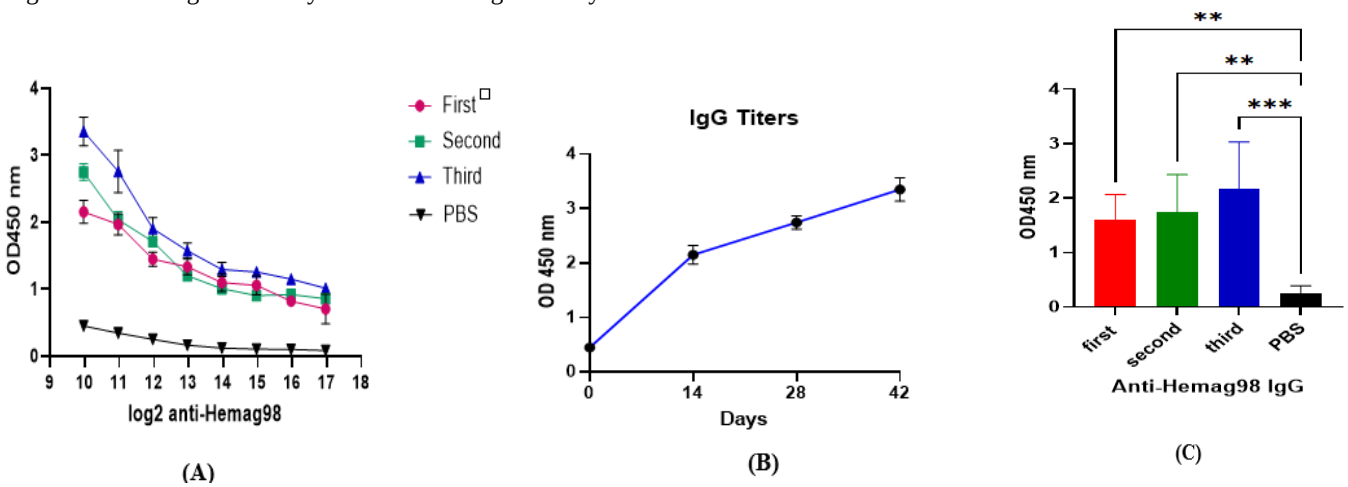

Fig 7. Measurement of Anti -Hemag98 specific IgG antibody titer in sera of vaccinated mice using ELISA. **A)** ELISA checkerboard to determine Hemag98 antibody titration. **B)** Anti-Hemag98 IgG raise after third injection. **C)** Assay the difference between test and control group using one way ANOVA analysis

[DOI: 10.52547/vacres.10.2.28]

DISCUSSION

The purpose of this study was to design and evaluate the immunogenicity of a multi-epitope vaccine against influenza disease. Compared to traditional vaccines, a multi-epitope vaccine candidate can provide broader protection against different strains of influenza virus due to the inclusion of multiple epitopes from various antigens [18]. Additionally, the use of bioinformatics tools and peptide synthesis techniques have allowed for the efficient identification and synthesis of the epitopes and rapid development of the vaccine candidates [19]. In the present study, bioinformatics tools were used to identify and select the most conserved and immunogenic epitopes of HA from various influenza strains. The selected epitopes were then combined into a single structure to create the multi-epitope vaccine candidate. This design allowed for the presentation of multiple antigens to the immune system in order to enhance the immune response, compared to traditional single-antigen vaccines. The vaccine candidate was expressed in a prokaryotic host to produce a recombinant protein, named Hemag98 which was then purified and administered to mice to assess its immunogenicity. The results demonstrated that the multi-epitope vaccine candidate induced a significant immune response in mice. The mice produced high levels of IgG antibodies, specific to the selected epitopes, indicating that the vaccine candidate was able to elicit an effective immune response.

We reviewed articles and used antigenic scoring to determine the target protein for designing the vaccine and found that the HA protein had the highest number of epitopes, as observed by Muñoz *et al*[20]. The distribution percentage of these epitopes for H1N1, H3N2, and H5N1 strains were determined to be 50.7%, 52.2%, and 67.7%, respectively. Moreover, based on the widespread prevalence of H1N1 and H3N2 strains and the pandemic potential of H5N1 and H7N9 strains, we selected these strains to identify protective, highly conserved, and antigenic epitopes for the vaccine design. The utilized online tools to select suitable humoral and cellular epitopes and evaluated various factors. Finally, 7 epitopes were chosen for the next steps of the study because they exhibited high protection percentages (ranging from 70% to 100%), high immunogenicity, and were non-toxic and non-allergenic.

HA protein is a type of glycoprotein consisting of two subunits of HA1 and HA2, which binds to sialic acid receptors on the surface of host cells and allows the virus to enter and initiate the infection. Furthermore, the HA protein is also the primary target for neutralizing antibodies produced by the immune system. These antibodies can block the attachment of the virus to host cells, preventing the infection[21]. Due to its abundance, compared to other surface proteins, HA is the primary target for developing subunit vaccines. Antibodies against this protein have high neutralizing properties and can prevent host infection by the virus (Skehel and Wiley 2000). These antibodies can easily be measured using the hemagglutination inhibition (HAI) test. The correlation between serum HAI titers and the protection provided by these antibodies is used to assess vaccine efficacy. In most cases, these neutralizing antibodies target variable regions of HA and therefore cannot serve as suitable epitopes for universal vaccine development. However, recently, monoclonal antibodies capable of cross-protection that can block or neutralize influenza virus infection have been identified in HA [22].

Another protein of interest for universal vaccine development is the M2 protein channel. This protein possesses a highly conserved external domain called M2e, which consists of 24 amino acids and can induce antibodies against a broad spectrum of influenza A viruses under both *in vitro* and *in vivo* conditions. The N-terminal of the M2e protein, consisting of 9 amino acids, is a highly conserved epitope (SLLTEVET) shared between the M1 and M2 proteins. Monoclonal antibodies against this epitope have been shown to inhibit the replication of H1 and H3 influenza viruses in MDCK cells [23].

Based on experiments and studies conducted, the M2e sequence is a suitable candidate for developing influenza vaccines. However, due to its small size, it results in transient immune responses. Various methods have been employed to improve the immunogenicity of M2e and to develop the best and the most effective vaccine against this virus. One approach includes the fusion of M2e with the tetramer protein GCN4, which is a eukaryotic transcriptional activator. As a result, the specific antibody responses to M2e are enhanced, and vaccinated mice are fully protected against influenza viruses [24]. Additionally, the M2e sequence has been fused with the core of the hepatitis B virus (M2e-HBC), which also provided protection against influenza viruses in mice [25]. Fusion of M2e with the N-terminal region of HSP70 from *Mycobacterium tuberculosis*, followed by expression and injection into animal models, demonstrates increased immunogenicity [26, 27]. Furthermore, the placement of multiple sequential copies of M2e, such as constructing four sequential copies of the M2e sequence and fusing it with TLR5 from *Salmonella typhimurium* as an adjuvant (STF2*4*M2e), has been performed. In the mouse model, the specific antibody response to M2e was increased, and the animals were protected against the virus [28]. These experiments have recently shown promising results in terms of immunogenicity and safety in human phase I and II clinical trials [28].

The ability to present foreign epitopes on the HA protein has made it a valuable tool in vaccine development. By presenting a foreign epitope to the HA protein, it can induce an immune response that targets both the HA protein and the foreign epitope. This allows for the development of vaccines that can protect against multiple strains of influenza or even other pathogens[8]. In this study, selected epitopes with appropriate arrangement were placed in the HA1 subunit of the pH1N1 influenza virus HA protein. The physicochemical analysis of the resulting structure revealed that it possesses adequate stability. Moreover, based on its aliphatic percentage (>50) and negative GRAVY score, it can be classified as a hydrophilic protein with strong thermal stability. The 3D structure of the modified HA1 subunit with epitopes was predicted using online software. This analysis confirmed that the epitopes introduced into the HA protein surface can be effectively recognized by epitope prediction tools, and they are accessible for immune responses.

The subsequent step in developing multi-epitope vaccine candidates involved producing the protein in a suitable host system. In this study, *E. coli* bacteria were chosen as the expression host due to its ease of handling, cost-effectiveness, and favorable production conditions. The investigation of expression in *E. coli* indicates that the expression of the Hemag98 protein began from the early hours after the induction and in the form of insoluble bodies. The level of expressed protein in the bacterial host did not show significant changes even after optimizing expression conditions such as temperature reduction after the induction, optimal IPTG concentration, and increased induction time. The maximum level recorded was 200 mg/L. Aguilar and colleagues also failed to improve expression of HA/H1N1 in *E. coli* BL21 DE3 and only achieved concentrations of 400 milligrams per liter again in the form of

[DOI: 10.52547/vacres.10.2.28]

insoluble bodies when using alternative strains such as *E. coli* Rosetta-gami and C41[29]. Sequence optimization including using bacterial codon preference and the use of specific tags seems to have an impact on enhancing protein solubility in bacteria. However, none of them were effective to reduce insoluble particles in this project. Although one key aspect in the production of recombinant protein is obtaining soluble forms and preventing the formation of insoluble particles, numerous studies have shown that proteins expressed as insoluble particles can achieve proper folding after extraction, denaturation, purification steps, and gradual removal of denaturing agents such as urea. In some cases, they can even perform better than soluble proteins in stimulating the immune system.

In this project, after the expression of bacteria in the form of insoluble particles, these particles were denatured using 8M urea. The protein was then purified using affinity chromatography with nickel particles, followed by dialysis to refold the protein and remove urea, considering its toxic properties in animal circulatory systems. To confirm the successful production and purity of the protein in the bacterial host, Western blot technique was employed. The presence of Hemag98 protein, with a molecular weight of 45 kDa, corresponding to the predicted molecular weight by the ProtParam tool, was observed on the nitrocellulose membrane during this analysis.

One of the strategies for investigating proper folding of recombinant proteins involves using ELISA technique with serum obtained from patients infected with the virus to measure the level of protein recognition by natural antibodies present in the serum. In this regard, serum from mice vaccinated with the complete H1N1 virus was used, and the results showed that the recombinant protein was recognized by IgG antibodies present in the serum (data not shown). In a study conducted by Aguilar and colleagues, the soluble form produced from the HA1 protein and the refolded form of this protein were examined for their ability to be detected by antibodies in the sera of individuals infected with the H1N1/2009 virus. The results demonstrated that the refolded form was recognized by the antibodies more than 90% better than the soluble form [29]. These results indicated that producing recombinant proteins in the form of inclusion bodies, with optimized refolding conditions, can be an effective approach for enhancing expression and mass production.

In the present study, the potency of a recombinant chimeric HA protein containing conserved epitopes of epidemic and pandemic potential strains was evaluated in a murine model. Results showed that the chimeric HA protein could potently stimulate humoral responses after even second immunization. The strong humoral responses would be due to the high antigenic properties of the designed construct which had been predicted by *in silico* tools (ANTIGENpro). Farsad, et al (2016) reported the properly folded HA1 could trigger the production of neutralizing antibodies and had protective activity in a rabbit model[30]. Khurana et al. (2011) described the presence of oligomerization region in N- terminal of HA1[31] and Verma et al. (2012) reported that bacterial produced H5 HA1could control viral loads and protect ferret from H5N1 challenges even better than H5N1 subunit vaccine, due to the oligomeric nature of HA1[32]. Our results showed that recombinant Hemag98 had proper folding as it recognized by anti-HA1 of PR8 immunized mice in its native conformation by ELISA (data not shown). This suggests that our recombinant protein might induce protection upon further animal studies in the future.

ACKNOWLEDGEMENTS

The authors would like to acknowledgement the Pasteur Institute of Iran for the financial support; Grant number: 1807.

CONFLICT OF INTEREST

The authors declare they have no conflict of interests.

REFERENCES

1. prevention Cfdca. 2022-2023 U.S. Flu Season: Preliminary In-Season Burden Estimates. 2022. https://www.cdc.gov/flu/about/burden/preliminary-in-season-estimates.htm.
2. Van de Sandt CE, Kreijtz JH, Rimmelzwaan GF. Evasion of influenza A viruses from innate and adaptive immune responses. Viruses. 2012;4(9):1438-76.
3. Wong S-S, Webby RJ. Traditional and new influenza vaccines. Clinical microbiology reviews. 2013;26(3):476-92.
4. Francis ME, King ML, Kelvin AA. Back to the Future for Influenza Preimmunity—Looking Back at Influenza Virus History to Infer the Outcome of Future Infections. Viruses. 2019;11(2):122.
5. Erbelding EJ, Post DJ, Stemmy EJ, Roberts PC, Augustine AD, Ferguson S et al. A universal influenza vaccine: the strategic plan for the National Institute of Allergy and Infectious Diseases. The Journal of infectious diseases. 2018;218(3):347-54.
6. Wang W-C, Sayedahmed EE, Sambhara S, Mittal SK. Progress towards the Development of a Universal Influenza Vaccine. Viruses. 2022;14(8):1684.
7. Li S, Schulman JL, Moran T, Bona C, Palese P. Influenza A virus transfectants with chimeric hemagglutinins containing epitopes from different subtypes. Journal of virology. 1992;66(1):399-404.
8. Li Z-N, Mueller SN, Ye L, Bu Z, Yang C, Ahmed R et al. Chimeric influenza virus hemagglutinin proteins containing large domains of the Bacillus anthracis protective antigen: protein characterization, incorporation into infectious influenza viruses, and antigenicity. Journal of virology. 2005;79(15):10003-12.
9. Geourjon C, Deléage G. SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. Comput Appl Biosci. 1995;11(6):681-4. doi:10.1093/bioinformatics/11.6.681.
10. Kim DE, Chivian D, Baker D. Protein structure prediction and analysis using the Robetta server. Nucleic Acids Res. 2004;32(Web Server issue):W526-31. doi:10.1093/nar/gkh468.
11. Wiederstein M, Sippl MJ. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res. 2007;35(Web Server issue):W407-10. doi:10.1093/nar/gkm290.
12. Gasteiger E, Hoogland C, Gattiker A, Duvaud Se, Wilkins MR, Appel RD et al. Protein identification and analysis tools on the ExPASy server. Springer; 2005.
13. Cheng J, Randall AZ, Sweredoski MJ, Baldi P. SCRATCH: a protein structure and structural feature prediction server. Nucleic acids research. 2005;33(suppl_2):W72-W6.
14. Saha S, Raghava GPS. AlgPred: prediction of allergenic proteins and mapping of IgE epitopes. Nucleic Acids Research. 2006;34(suppl_2):W202-W9. doi:10.1093/nar/gkl343.
15. Avalós PP. 1. OPTIMIZER: a web server for optimizing the codon usage of DNA sequences. Pere Puigbò, Eduard Guzmán, Antoni. Codon usage adaptation in prokaryotic genomes.39.
16. Chang AY, Chau V, Landas JA, Pang Y. Preparation of calcium competent Escherichia coli and heat-shock transformation. JEMI methods. 2017;1:22-5.
17. Khosravi Node F, Behzadian F, Mazaheri V, Shokouhi H, Saleh M, Farahmand B. Purification of Influenza virus A (H1N1) recombinant Hemagglutinin (HA1) and polyclonal antibody production. Tehran University Medical Journal TUMS Publications. 2017;75(8):585-92.
18. Romeli S, Hassan SS, Yap WB. Multi-epitope peptide-based and vaccinia-based universal influenza vaccine candidates subjected to clinical trials. The Malaysian journal of medical sciences: MJMS. 2020;27(2):10.
19. Burton DR. What are the most powerful immunogen design vaccine strategies? Reverse vaccinology 2.0 shows great promise. Cold Spring Harbor perspectives in biology. 2017;9(11):a030262.

[DOI: 10.52547/vacres.10.2.28]

20. Muñoz-Medina JE, Sánchez-Vallejo CJ, Méndez-Tenorio A, Monroy-Muñoz IE, Angeles-Martínez J, Santos Coy-Arechavaleta A et al. In silico identification of highly conserved epitopes of influenza A H1N1, H2N2, H3N2, and H5N1 with diagnostic and vaccination potential. BioMed Research International. 2015;2015.
21. Copeland CS, Doms RW, Bolzau EM, Webster RG, Helenius A. Assembly of influenza hemagglutinin trimers and its role in intracellular transport. Journal of Cell Biology. 1986;103(4):1179-91.
22. Martinez O, Tsibane T, Basler CF. Neutralizing anti-influenza virus monoclonal antibodies: therapeutics and tools for discovery. International reviews of immunology. 2009;28(1-2):69-92.
23. Gottlieb T, Ben-Yedidia T. Epitope-based approaches to a universal influenza vaccine. Journal of autoimmunity. 2014;54:15-20.
24. De Filette M, Martens W, Roose K, Deroo T, Vervalle F, Bentahir M et al. An influenza A vaccine based on tetrameric ectodomain of matrix protein 2. J Biol Chem. 2008;283(17):11382-7. doi:10.1074/jbc.M800650200.
25. De Filette M, Martens W, Smet A, Schotsaert M, Birkett A, Londoño-Arcila P et al. Universal influenza A M2e-HBc vaccine protects against disease even in the presence of pre-existing anti-HBc antibodies. Vaccine. 2008;26(51):6503-7.
26. Ebrahimi SM, Tebianian M. Heterologous expression, purification and characterization of the influenza A virus M2e gene fused to Mycobacterium tuberculosis HSP70 359–610 in prokaryotic system as a fusion protein. Molecular biology reports. 2010;37:2877-83.
27. Attaran H, Nili H, Tebianian M. Immunogenicity and protective efficacy of recombinant M2e. Hsp70c (Hsp70 359–610) fusion protein against influenza virus infection in mice. Virologica Sinica. 2014;29:218-27.
28. Turley CB, Rupp RE, Johnson C, Taylor DN, Wolfson J, Tussey L et al. Safety and immunogenicity of a recombinant M2e–flagellin influenza vaccine (STF2. 4xM2e) in healthy adults. Vaccine. 2011;29(32):5145-52.
29. Aguilar-Yáñez JM, Portillo-Lara R, Mendoza-Ochoa GI, García-Echauri SA, López-Pacheco F, Bulnes-Abundis D et al. An Influenza A/H1N1/2009 Hemagglutinin Vaccine Produced in Escherichia coli. PLOS ONE. 2010;5(7):e11694. doi:10.1371/journal.pone.0011694.
30. Farsad A, Malekzadeh-Shafaroudi S, Moshtaghi N, Fotouhi F, Zibaee S. Expression of HA1 antigen of H5N1 influenza virus as a potent candidate for vaccine in bacterial system. Iranian Journal of Veterinary Research. 2016;17(4):237.
31. Khurana S, Verma S, Verma N, Crevar CJ, Carter DM, Manischewitz J et al. Bacterial HA1 vaccine against pandemic H5N1 influenza virus: evidence of oligomerization, hemagglutination, and cross-protective immunity in ferrets. Journal of virology. 2011;85(3):1246-56.
32. Verma S, Dimitrova M, Munjal A, Fontana J, Crevar CJ, Carter DM et al. Oligomeric recombinant H5 HA1 vaccine produced in bacteria protects ferrets from homologous and heterologous wild-type H5N1 influenza challenge and controls viral loads better than subunit H5N1 vaccine by eliciting high-affinity antibodies. Journal of virology. 2012;86(22):12283-93.

[DOI: 10.52547/vacres.10.2.28]