CME LIVE OPEN ACCESS

Vaccine and Immunization Research and Development

Research Article

Computational Approaches to Influenza Vaccine Development

Yukti Sabikhi* and Anshika Singh

School of Engineering Science and Technology, Jamia Hamdard University, New Delhi 110062, India

Corresponding Author: Yukti Sabikhi, School of Engineering Science and Technology, Jamia Hamdard University, New Delhi 110062, India.

Received: 2026 Jan 05 **Accepted:** 2026 Jan 25 **Published:** 2026 Feb 03

Abstract

Influenza is an extremely contagious respiratory illness that remains a significant public health issue because of its swift antigenic drift and shift, resulting in the appearance of new strains. Seasonal vaccines necessitate regular updates and frequently exhibit lower effectiveness, emphasizing the demand for universal and more potent vaccine options. This research utilized computational methods to evaluate the Hemagglutinin (HA) Chain A protein of the influenza virus as a possible vaccine target, due to its essential function in viral attachment, membrane fusion, and immune recognition. The HA protein sequence was sourced from the NCBI database and examined using ProtParam to assess physicochemical attributes, PSIPRED for predicting secondary structures, and DiANNA for evaluating disulfide bonds. VaxiJen was utilized to assess antigenicity, and AllerTOP was employed to confirm non-allergenic potential. 3D structural modeling with high confidence was conducted using trROSETTA to confirm structural integrity and find possible epitopic areas. Findings suggested that the HA protein has advantageous stability characteristics, an elevated antigenicity score, and is expected to be non-allergenic, exhibiting a strong tertiary structure reinforced by disulfide bonds. The trROSETTA model reached a TM-score of 0.908, validating precise folding and epitope availability. These results illustrate the promise of HA Chain A as a viable vaccine candidate and highlight the importance of combined in silico analyses for speeding up vaccine development. This computational pipeline provides a systematic approach for the logical creation of next-generation influenza vaccines, facilitating experimental validation and subsequent clinical application.

Keywords: Influenza, Hemagglutinin, Vaccine Development, In Silico Analysis, Antigenicity Prediction, Allergenicity Prediction, Structural Modeling, ProtParam, PSIPRED, DiANNA, VaxiJen, AllerTOP, trROSETTA and Computational Biology, Universal Influenza Vaccine

1. Introduction

Influenza is a highly contagious respiratory disease that can infect humans and is caused by a variety of RNA influenza viruses. Influenza-related complications can result in serious morbidity and death [1]. Up to 500,000 people worldwide lose their lives to influenza-related complications each year. According to the Center for Disease Control and Prevention (CDC), among Americans, seasonal influenza infection has caused between 9.3 and 45 million infections, 140,000 and 810,000 hospitalizations, and 12,000 and 61,000 fatalities every year since 2010 [2]. For high-risk people or groups, influenza infection complications can be serious or even fatal [3]. Fever, chills, headache, weakness, red eyes, sore throat, dry cough, and nasal discharge are all signs of influenza [4]. Influenza viruses frequently change their antigens, which leads to rapid evolution. The processes by which the virus mutates and produces new strains are referred to as antigenic drift and shift [5]. The HA (Hemagglutinin), NA (Nucleoprotein), and M2(Matrix protein) are found within the outer envelope of the viral particle, with HA and NA being prominently displayed, making them significant immunogenic proteins. During the infection process, HA is crucial for viral entry, including the stages of attachment and penetration. HA serves as the primary antigenic glycoprotein on the surface of the influenza virus, capable of triggering neutralizing antibody responses in the host [6-9]. NA plays a role in severing the terminal sialic acid residues on cell surfaces, thereby assisting in the release of progeny virions from infected cells and is key in the dissemination of influenza viruses [10]. Influenza viruses are categorized into 18 hemagglutinin (HA) subtypes (H1-H18) and 11 neuraminidase (NA) subtypes (N1-N11), due to the antigenic variability of HA and NA [11-13].

Hemagglutinin is the primary protein found on the surface of the influenza virus and plays a crucial role in the viral entry process through its binding to receptors and facilitating membrane fusion [14]. As the main antigenic protein, HA demonstrates the highest levels of adaptive evolution compared to other influenza proteins [15]. Despite significant sequence diversity among the various subtypes, the HA protein retains essential components

such as the cleavage site, secretory signal, fusion domain, transmembrane domain, and cytoplasmic tail, along with common structural motifs [16]. Vaccination continues to be the most effective method for preventing influenza and its associated complications. Seasonal vaccines are updated each year to align with circulating strains and demonstrate an effectiveness of 40-60%; however, they often do not reach the WHO target of 75% coverage, largely due to vaccine hesitancy, restricted access to healthcare, and spreading misinformation. The two primary types of vaccines are live attenuated influenza vaccines (LAIV), which are administered as nasal sprays to healthy adults aged 2 to 49, and inactivated influenza vaccines (IIV), which are advised for most age groups, including those at higher risk [17,18]. Emerging alternatives such as recombinant hemagglutinin (Flublok) and cell-based vaccines enhance effectiveness by eliminating egg-adapted mutations. Antiviral medications encompass neuraminidase inhibitors (oseltamivir, zanamivir) and cap-dependent endonuclease inhibitors (baloxavir marboxil), which yield the best results when given promptly [19]. The growing issue of antiviral resistance and inconsistent vaccine efficacy caused by antigenic drift highlight the necessity for universal vaccines and advanced therapeutic options [20,21]. In this research, the Hemagglutinin protein (Chain A) of the influenza virus was chosen as the target for vaccine development because of its essential function in viral entry and immune detection. The FASTA sequence underwent analysis through various computational tools, such as ProtParam for physicochemical properties, PSIPRED for predicting secondary structure, DiANNA for analyzing disulfide bonds, VaxiJen for predicting antigenicity, AllerTOP for assessing allergenicity, and trROSETTA for structural modeling. These analyses facilitated the assessment of various parameters to determine if this protein might act as a potential vaccine candidate, establishing a solid basis for additional in silico vaccine development.

2. Methodology

For the development of a possible vaccine against influenza, the Hemagglutinin Chain A protein was chosen as the target because of its essential function in viral entry and immune detection. The approach included a sequential computational examination (Fig.1) , as outlined below:

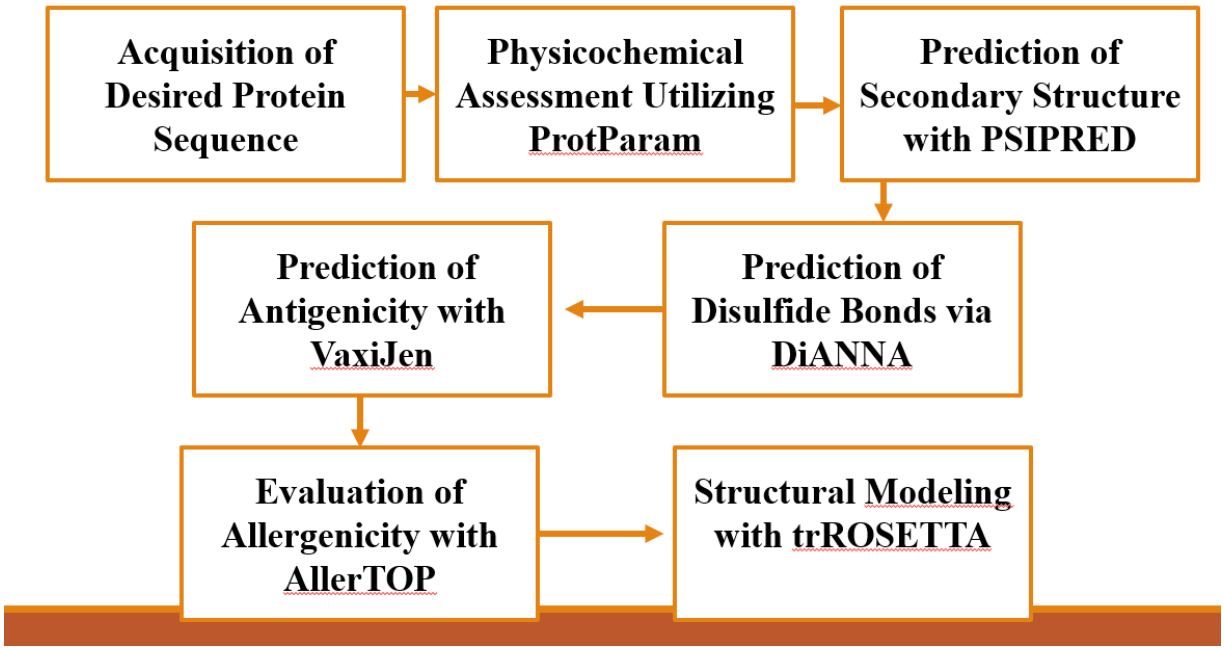

Figure 1: Workflow of Vaccine Design

2.1. Acquisition of Desired Protein Sequence

The FASTA sequence for Hemagglutinin Chain A was obtained from the NCBI protein database [22]. Its accession ID and locus details were noted to guarantee precise identification. This sequence acted as the basis for all later computational analyses.

2.2. Physicochemical Assessment Utilizing ProtParam

The obtained sequence was analyzed using ProtParam to determine physicochemical properties including molecular weight, theoretical isoelectric point (pI), amino acid composition, instability index, aliphatic index, and GRAVY. These parameters assist in assessing the stability, solubility, and appropriateness of the protein as a potential vaccine candidate [23].

2.3. Prediction of Secondary Structure with PSIPRED

The sequence was examined with PSIPRED a dependable tool for forecasting secondary structural components such as alpha-helices, beta-sheets, and coils [24]. This phase offered understanding of the structural arrangement and possible antigenic areas.

2.4. Prediction of Disulfide Bonds Via DiANNA

The protein sequence was sent to DiANNA, utilizing neural networks and SVM-based algorithms to forecast cysteine

connectivity and the creation of disulfide bonds. The existence of disulfide bonds signifies the stability of tertiary structure, an essential factor for vaccine development [25].

2.5. Prediction of Antigenicity with VaxiJen

The antigenic capacity of Hemagglutinin Chain A was assessed using VaxiJen v2.0, applying a threshold of 0.4 for viral proteins. A score exceeding this threshold indicates that the protein may function as a likely vaccine candidate [26].

2.6. Evaluation of Allergenicity with AllerTOP

To ensure safety, the protein was evaluated for allergenicity through AllerTOP v2.0, which forecasts if a protein is an allergen or non-allergen based on sequence-derived descriptors. Only proteins that do not cause allergies were taken into account for the creation of the vaccine [27].

2.7. Structural Modeling with TROSETTA

Ultimately, TROSETTA was utilized to forecast the 3D tertiary structure of Hemagglutinin Chain A. This modeling confirmed the predictions of secondary and disulfide bonds and allowed for the identification of epitopic regions appropriate for vaccine targeting [28].

3. Result

3.1. Retrieval of Target Protein Sequence

The desired protein Hemagglutinin Chain A, was obtained from the NCBI database in FASTA format. Its accession ID and locus details (Table-1) were documented to guarantee accurate identification. This sequence acted as the core data for all later computational evaluations in the vaccine design process.

Protein	Accesion ID	Locus	PDB
Chain A, Hemagglutinin	8TP2_A	8TP2_A	8TP2

Table 1: Desired Protein Details from Ncbi

3.2. FASTA Sequence–

MAIIYLILLFTAVRGDQICIGYHANNSTEKVDTILERNVT-
VTHAKDILEKTHNGKLCKLNGIPPLELGDCSIAGWLLGN-
PECDRLLSVPEWSYIMEKENPRDGLCYPGSFNDYEELKH-
LLSSVKHFEKVKILPKDRWTQHTTTGGSRACAVSGNPSF-
FRNMVWLTKKGSNYPVAKGSYNNTSGEQMLIIWGVHH-
PNDETEQRTLYQNVGTYVSVGTSTLNKRSTPDIATRPKVNGQ-
GGRMEFSWTLLDMWDTINFESTGNLIAPEYGFKISKRGSS-
GIMKTEGTLENCETKCQTPLGAINTTLPFHNVHPLTIGECP-
KYVKSEKLVLATGLRNVPQIESRGLFGAIAGFIEGGWQGM-
VDGWYGYHHSNDQGSGYAADKESTQKAFDGITNKVNS-
VIEKMNTQFEAVGKEFSNLERRLENLNKKMEDGFLDVW-
TYNAELLVLMENERTLDFHDSNVKNLYDKVRMQLRDNVKEL-
GNGCFEFYHKCDDECMNSVKNGTYDYPKYEEES

3.3. Physicochemical Analysis Using ProtParam

The obtained FASTA sequence of Hemagglutinin Chain A was entered into ProtParam to evaluate its physicochemical characteristics (Table-2) such as molecular weight, predicted isoelectric point (pI), amino acid composition, instability index, aliphatic index, and grand average of hydropathicity (GRAVY). These parameters aided in evaluating the stability, solubility, and appropriateness of the protein as a possible vaccine candidate.

Parameters	Result
Number of amino acids:	506
Theoretical pI:	5.84
Molecular weight:	57066.40
Total number of negatively charged residues (Asp + Glu):	65
Total number of positively charged residues (Arg + Lys):	55
Total number of atoms:	7916
Instability index:	31.89
Aliphatic index:	73.56
Grand average of hydropathicity (GRAVY):	-0.532

Table 2: Physicochemical Characteristics of Protein

3.4. Secondary Structure Prediction Using PSIPRED

The FASTA sequence of Hemagglutinin Chain A was uploaded to PSIPRED for predicting elements of its secondary structure, such as alpha-helices, beta-sheets, and coils. This analysis revealed details about the protein's structure, critical for pinpointing possible antigenic areas for vaccine development. Psipred, Memsat and aatype are demonstrated in **Fig.2, Fig.3** and **Fig.4**.

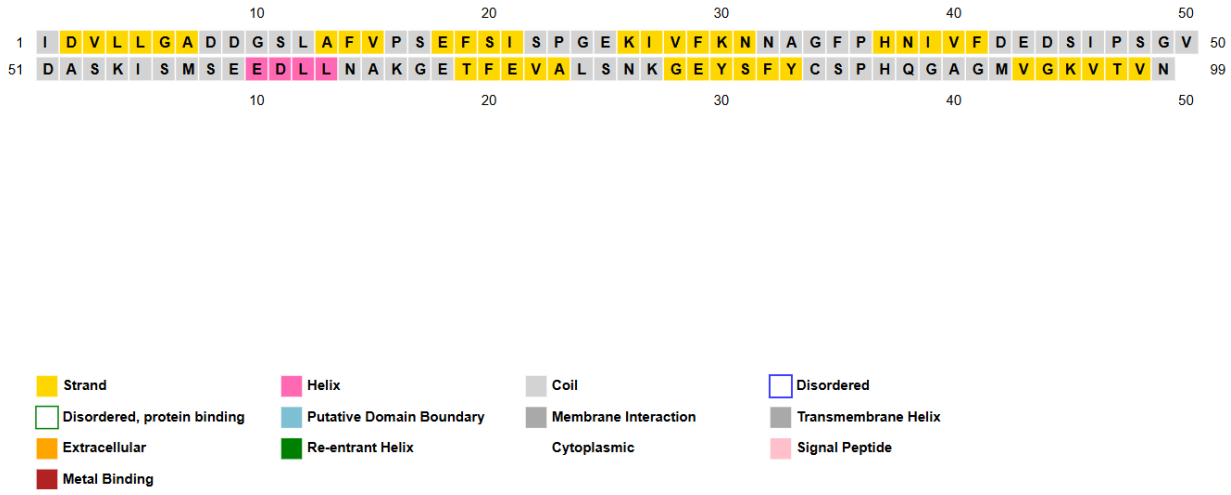

Figure 2: Psipred

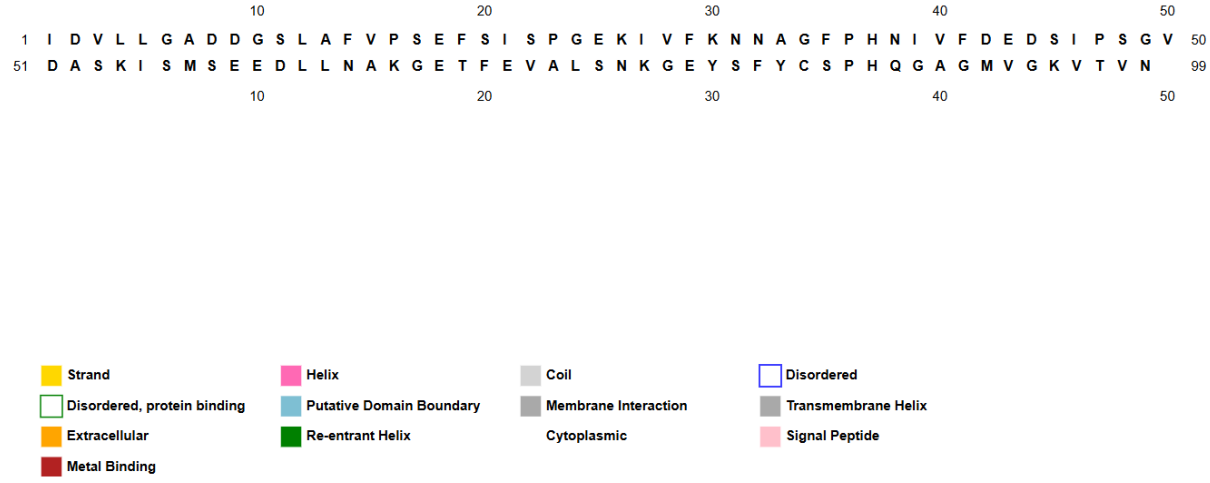

Figure 3: Memsat

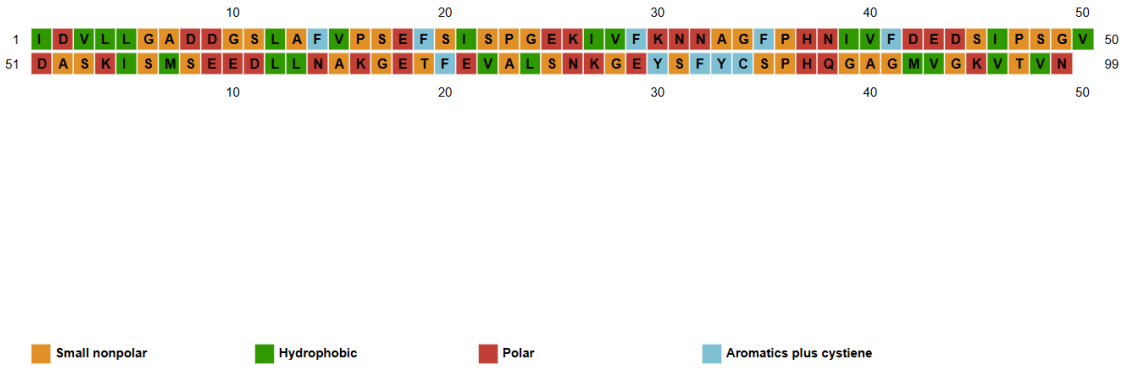

Figure 4: Aatypes

3.5. Disulfide Bond Prediction Using DiANNA

The FASTA sequence of Hemagglutinin Chain A was uploaded to DiANNA to determine the existence and location of disulfide bonds in the protein. Recognizing these bonds is crucial for comprehending the protein's structural integrity and for identifying areas appropriate for vaccine candidate development. The predicted **disulfide bonds** indicate that Hemagglutinin Chain A has a **stable tertiary structure**, making it a promising candidate for **vaccine design**.

3.6. Antigenicity Prediction Using VaxiJen

The sequence of Hemagglutinin Chain A was evaluated with VaxiJen v2.0 at a cutoff of 0.4 to forecast antigenic capability (Table-3) . Proteins that exceed the threshold are seen as likely vaccine candidates.

Threshold for this model:	Overall Prediction for the Protective Antigen	Result
0.4	**0.4612**	Probable **ANTIGEN**

Table 3: Antigenicity Prediction

3.7. Allergenicity Assessment Using AllerTOP

The protein underwent analysis using AllerTOP v2.0 to evaluate its allergenic potential, confirming that the candidate vaccine is non-allergenic and suitable for subsequent development.

- Most similar protein: sp|Q9H799|CE042_HUMAN Uncharacterized protein C5orf42 OS=Homo sapiens GN=C5orf42 PE=1 SV=4
- Classification based on the most similar protein: Probable NON-ALLERGEN

3.8. Structural Modeling Using TrROSETTA

Ultimately, the protein arrangement was simulated with TROSETTA to illustrate its three-dimensional shape, confirm secondary and tertiary structure forecasts, and locate possible epitopic areas for vaccine development. The estimated TM-score of the model is extremely high at 0.908. The TM-score serves as a measure for assessing the similarity of two protein structures, ranging from 0 to 1. A score exceeding 0.5 suggests a model with the correct overall fold or topology, while a score nearing 1 represents a highly precise model. The graph (Fig. 5) displays the Predicted Local Distance Difference Test (LDDT) scores for each amino acid position within the protein. The majority of the scores fall within the range of 70 to 90, indicating that the local structure of the protein is anticipated with great precision. The confidence of the model is very high. The picture shows the anticipated three-dimensional form of the protein (Fig. 6).

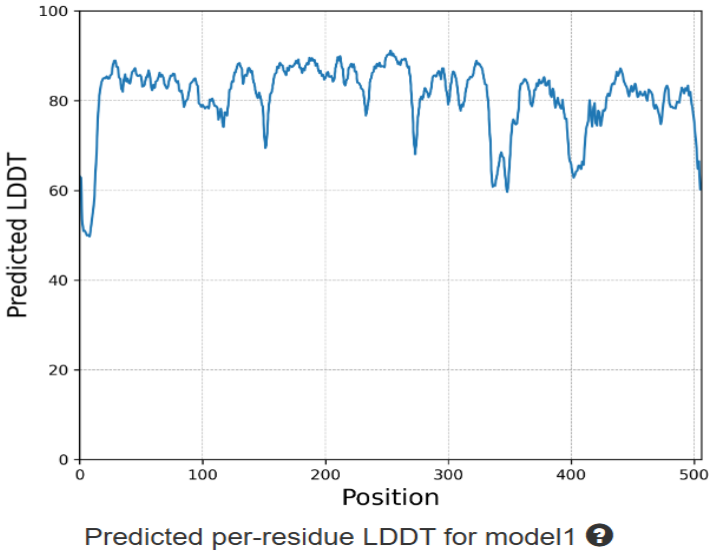

Figure 5: LDDT Graph

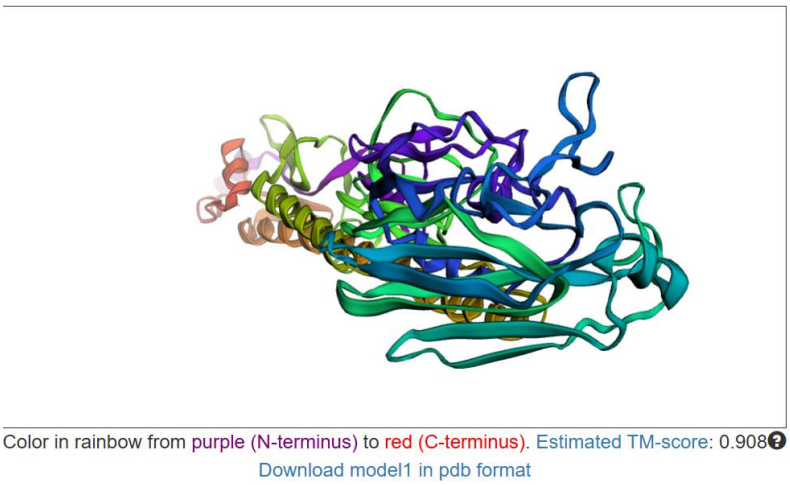

Figure 6: Protein Structure

According to the trRosetta findings you shared, the model depicts the anticipated 3D configuration of a protein along with its quality. The model's confidence is characterized as "extremely high." The elevated confidence score indicates that the forecasted 3D structure is probably correct and trustworthy. The model was created using deep learning techniques along with homologous templates, which are similar protein structures that already exist. The findings additionally offer an expected distribution (Fig. 7) of inter-residue distances and orientations, beneficial for the subsequent examination of the protein's folded configuration. This result reflects a robust and assured forecast of the protein's three-dimensional structure.

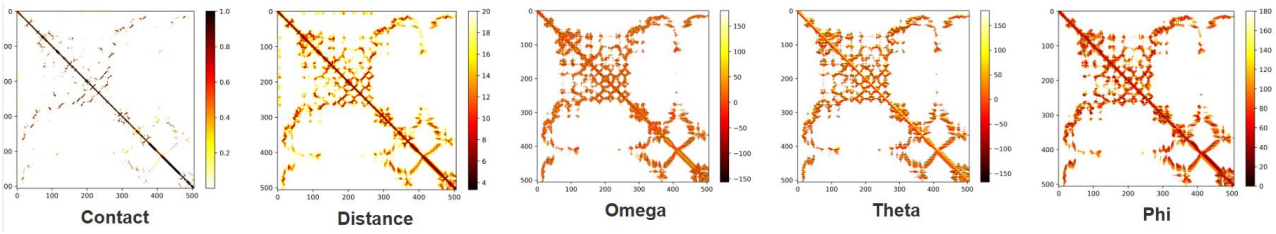

Figure 7: Predicted Inter-Residue Distances and Orientations

4. Conclusion

The research utilized an extensive computational method to assess the Hemagglutinin Chain A protein of the influenza virus as a possible vaccine candidate. Utilizing various bioinformatics assessments, the protein exhibited advantageous traits, such as reliable physicochemical properties, a clearly defined secondary and tertiary structure, along with disulfide bonds suggesting structural stability. The prediction of antigenicity validated its ability to induce an immune response, while evaluations of allergenicity guaranteed its safety for human application. The high-quality 3D structural model produced by trROSETTA further confirmed the protein's appropriateness for vaccine development by pinpointing possible epitopic areas. These results underscore the potential of Hemagglutinin Chain A as a focus for developing influenza vaccines. The computational approaches employed in this research establish a strong foundation for in silico vaccine creation, setting the stage for forthcoming experimental confirmation and clinical testing. Utilizing these advanced tools, researchers can expedite the creation of more effective and universal influenza vaccines, tackling the issues presented by antigenic drift and shift, and ultimately diminishing the worldwide impact of influenza-related illness and death.

References

1. Cummings, C. N., O'Halloran, A. C., Azenkot, T., Reingold, A., Alden, N. B., Meek, J. I., ... & Garg, S. (2022). Hospital-acquired influenza in the United States, FluSurv-NET, 2011–2012 through 2018–2019. *Infection Control & Hospital Epidemiology, 43*(10), 1447-1453.
2. Gaitonde, D. Y., Moore, F. C., & Morgan, M. K. (2019). Influenza: diagnosis and treatment. *American family physician, 100*(12), 751-758.
3. Clohisey, S., & Baillie, J. K. (2019). Host susceptibility to severe influenza A virus infection. *Critical care, 23*(1), 303.
4. Nakagawa, H., Noma, H., Kotake, O., Motohashi, R., Yasuda, K., & Shimura, M. (2017). Optic neuritis and acute anterior uveitis associated with influenza A infection: a case report. *International medical case reports journal*, 1-5.
5. Kim, H., Webster, R. G., & Webby, R. J. (2018). Influenza virus: dealing with a drifting and shifting pathogen. *Viral immunology, 31*(2), 174-183.
6. Chromikova, V., Tan, J., Aslam, S., Rajabhathor, A., Bermudez-Gonzalez, M., Ayllon, J., ... & Krammer, F. (2020). Activity of human serum antibodies in an influenza virus hemagglutinin stalk-based ADCC reporter assay correlates with activity in a CD107a degranulation assay. *Vaccine, 38*(8), 1953-1961.
7. Krammer, F., & Palese, P. (2019). Universal influenza virus vaccines that target the conserved hemagglutinin stalk and conserved sites in the head domain. *The Journal of infectious diseases, 219*(Supplement_1), S62-S67.
8. Sriwilaijaroen, N., & Suzuki, Y. (2012). Molecular basis of the structure and function of H1 hemagglutinin of influenza virus. *Proceedings of the Japan Academy, Series B, 88*(6), 226-249.
9. Kosik, I., & Yewdell, J. W. (2019). Influenza hemagglutinin and neuraminidase: Yin–Yang proteins coevolving to thwart immunity. *Viruses, 11*(4), 346.
10. Jagadesh, A., Salam, A. A. A., Mudgal, P. P., & Arunkumar, G. (2016). Influenza virus neuraminidase (NA): a target for antivirals and vaccines. *Archives of virology, 161*(8), 2087-2094.
11. Tong, S., Zhu, X., Li, Y., Shi, M., Zhang, J., Bourgeois, M., ... & Donis, R. O. (2013). New world bats harbor diverse influenza A viruses. *PLoS pathogens, 9*(10), e1003657.
12. Wohlbold, T. J., & Krammer, F. (2014). In the shadow of hemagglutinin: a growing interest in influenza viral neuraminidase and its role as a vaccine antigen. *Viruses, 6*(6), 2465-2494.
13. Gorka, M., Schinköthe, J., Ulrich, R., Ciminski, K., Schwemmle, M., Beer, M., & Hoffmann, D. (2020). Characterization of experimental oro-nasal inoculation of Seba's short-tailed bats (Carollia perspicillata) with bat influenza A virus H18N11. *Viruses, 12*(2), 232.
14. Fields, B.N., 2007. *Fields' virology* (Vol. 1). Lippincott Williams & Wilkins.
15. Bhatt, S., Holmes, E. C., & Pybus, O. G. (2011). The genomic rate of molecular adaptation of the human

influenza A virus. *Molecular biology and evolution, 28*(9), 2443-2451.
16. Nobusawa, E., Aoyama, T., Kato, H., Suzuki, Y., Tateno, Y., & Nakajima, K. (1991). Comparison of complete amino acid sequences and receptor-binding properties among 13 serotypes of hemagglutinins of influenza A viruses. *Virology, 182*(2), 475-485.
17. Grohskopf, L. A. (2019). Prevention and control of seasonal influenza with vaccines: recommendations of the Advisory Committee on Immunization Practices—United States, 2019–20 influenza season. *MMWR. Recommendations and Reports, 68*.
18. "Recommendations for prevention and control of influenza in children, 2024–2025: policy statement." *Pediatrics* 154, no. 4 (2024): e2024068507.
19. Dunkle, L. M., Izikson, R., Patriarca, P., Goldenthal, K. L., Muse, D., Callahan, J., & Cox, M. M. (2017). Efficacy of recombinant influenza vaccine in adults 50 years of age or older. *New England Journal of Medicine, 376*(25), 2427-2436.
20. Hayden, F. G., Sugaya, N., Hirotsu, N., Lee, N., de Jong, M. D., Hurt, A. C., ... & Watanabe, A. (2018). Baloxavir marboxil for uncomplicated influenza in adults and adolescents. *New England Journal of Medicine, 379*(10), 913-923.
21. Ikematsu, H., Hayden, F. G., Kawaguchi, K., Kinoshita, M., de Jong, M. D., Lee, N., ... & Uehara, T. (2020). Baloxavir marboxil for prophylaxis against influenza in household contacts. *New England Journal of Medicine, 383*(4), 309-320.
22. Schoch, C. L., Ciufo, S., Domrachev, M., Hotton, C. L., Kannan, S., Khovanskaya, R., ... & Karsch-Mizrachi, I. (2020). NCBI Taxonomy: a comprehensive update on curation, resources and tools. *Database, 2020*, baaa062.
23. Gasteiger, E., Jung, E., & Bairoch, A. (2001). SWISS-PROT: connecting biomolecular knowledge via a protein database. *Current issues in molecular biology, 3*(3), 47-55.
24. Buchan, D. W., & Jones, D. T. (2019). The PSIPRED protein analysis workbench: 20 years on. *Nucleic acids research, 47*(W1), W402-W407.
25. Ferrè, F., & Clote, P. (2005). DiANNA: a web server for disulfide connectivity prediction. *Nucleic acids research, 33*(suppl_2), W230-W232.
26. Doytchinova, I. A., & Flower, D. R. (2007). VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC bioinformatics, 8*(1), 4.
27. Dimitrov, I., Bangov, I., Flower, D. R., & Doytchinova, I. (2014). AllerTOP v. 2—a server for in silico prediction of allergens. *Journal of molecular modeling, 20*(6), 2278.
28. Rohl, C. A. (2005). Protein structure estimation from minimal restraints using Rosetta. In *Methods in enzymology* (Vol. 394, pp. 244-260). Academic Press.