

Multi-epitope peptide sequence in-silico construction from HGV genome

Kumar Sharp (✉ ksharp0016@gmail.com)

Government Medical College and Hospital, Jalgaon <https://orcid.org/0000-0002-5034-8326>

Research Article

Keywords: Hepatitis G, vaccine, hepatitis C, HIV, peptide, reverse vaccinology

Posted Date: August 6th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-53866/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

In this study I have approached through in-silico method or reverse vaccinology taking advantage of the genome sequence of hepatitis G virus. It serves its benefit of identifying antigens seen by both conventional as well as discovering any novel antigen. This peptide candidate can serve a triple purpose of hepatitis C vaccine, hepatitis G vaccine and HIV management addition. 89.2% of the residues were in the favoured region of Ramachandran plot. These points make it favourable for in-vitro trials and further refinement. Because of the high similarity of hepatitis C genome to hepatitis G genome, it is highly probable that this peptide sequence might act as both hepatitis C and hepatitis G vaccine. Patients with past or current HGV infection have higher CD4+ lymphocyte counts and better AIDS-free survival rates. This peptide sequence might cause a breakthrough in the treatment of HIV without exposing them to develop hepatitis.

Introduction

In this study I have approached through in-silico method or reverse vaccinology taking advantage of the genome sequence of hepatitis G virus. It serves its benefit of identifying antigens seen by both conventional as well as discovering any novel antigen [1]. With technological advancement in the field of immunology these studies have become easier and more accurate [2]. This peptide candidate can serve a triple purpose of hepatitis C vaccine, hepatitis G vaccine and HIV management addition.

Methodology

The procedure used in this study is entirely based on two previous studies [3][4]. It will not be repeated here but is summarised below in figure 1 after obtaining necessary permission from its authors.

The modifications done in this study is as follows:

The genome sequence of hepatitis G virus was taken from NCBI virus database [5] (accession number:NC001710) Polyprotein precursor protein of this genome is used in this study. Docking of final peptide sequence was done with tertiary structure of toll-like receptor 7 obtained from Protein Data Bank [6] (PDB ID:5GMF).

Result

The final multi-epitope sequence formed after performing step 1 and 2 of methodology comprised of 214 amino acids:

```
GIINTLQKYYCRVRGGRCVLSCLPKEEQIGKCSTRGRKCCRRKEAAAKAVEAGVT  
WYAAYLLDFVFLLAAYVTDAAAIQAAYDVALETELYGPGPGWPLYQAGLAVRP  
GKSGPGPGAASYLMGLGVGGNAQGP GPGPLYQAGLAVRPGKSAGPGPGAVFFSGL  
APLRMHPDGP GPGASYLMGLGVGGNAQTGP GPGVFFSGLAPLRMHPDV
```

The above sequence distribution is as follows:

GIINTLQKYYCRVRGGRCVLSCLPKKEEQIGKCSTRGRKCCRRK EAAAK

(The above sequence is beta-defensin adjuvant sequence with EAAAK linker)

AVEAGVTWY AAY LLDFVFLVLL AAY VTDAVAAIQ AAY DVALETELY GPGPG

(Polyprotein precursor CTL epitopes linked to each other by AAY linker with highest immunogenicity and at the end with GPGPG linker to HTL epitopes)

| | | | |
|-----------------|-------|-----------------|-------|
| WPLYQAGLAVRPGKS | GPGPG | AASYLMGLGVGGNAQ | GPGPG |
| PLYQAGLAVRPGKSA | GPGPG | AVFFSGLAPLRMHPD | GPGPG |

ASYLMGLGVGGNAQT GPGPG VFFSGLAPLRMHPDV

(Polyprotein precursor HTL epitope with GPGPG linker)

The antigenicity prediction as in step 3 by Vaxijen server predicted it to be a probable antigen with score 0.462. The allergenicity prediction as in step 3 by Allpred server predicted it to be a non-allergen with score of -0.86908598 (positive predictive value is 0% and negative predictive value is 0%).

Physio-chemical properties as estimated by step 4 using ProtParam server gave the following results:

Molecular weight:21806.21 Daltons

Theoretical pI=9.28

Estimated half-life in E. coli: >10 hours (in vivo)

Instability index: 27.48(stable)

Aliphatic index: 80.79

Grand average of hydropathicity (GRAVY): 0.096

The secondary structure of the final multi-epitope sequence was computed using PHYRE 2 server: 40% comprised of alpha-helix, 3% of beta-strand, 13% of transmembrane helix and 22% was disordered (Figure 2).

The tertiary structure obtained from PHYRE2 server was subjected to refinement by Galaxy Refine tool which generated 5 models as follows (Table 1):

| Model | GDT-HA | RMSD | MolProbity | Clash Score | Poor rotamers | Rama favoured |
|---------|--------|-------|------------|-------------|---------------|---------------|
| Initial | 1.0000 | 0.000 | 4.192 | 141.9 | 10.2 | 73.1 |
| Model 1 | 0.9276 | 0.473 | 2.239 | 12.1 | 0.7 | 85.8 |
| Model 2 | 0.9241 | 0.487 | 2.281 | 11.8 | 1.4 | 88.2 |
| Model 3 | 0.9241 | 0.483 | 2.092 | 8.9 | 0.0 | 87.3 |
| Model 4 | 0.9287 | 0.461 | 2.167 | 12.1 | 0.0 | 89.2 |
| Model 5 | 0.9077 | 0.499 | 2.231 | 13.1 | 0.7 | 87.7 |

Table 1: Galaxy Refine structure models

Model 4 was chosen as the best tertiary structure of the sequence for further analysis. It was visualized using UCSF Chimera software [7] (Figure 3).

Ramachandran plot analysis by RAMPAGE server (Figure 4) gave the following result: Number of residues in favoured region (~98.0% expected): 189 (89.2%); Number of residues in allowed region (~2.0% expected): 15 (7.1%); Number of residues in outlier region: 8 (3.8%)

The predicted B-cell linear epitopes were calculated using Ellipro suite (Figure 5 and 6).

Toll-like receptor 7 was docked with the final model by PatchDock server and top 10 results were refined using FireDock server. Solution number 2 was the most favourable binding conformation with global energy at -4.41 and 0.00 repulsive Vander Waal forces. The docked model was visualized using UCSF Chimera (Figure 7).

Discussion

The protein sequence is predicted to be antigenic as well as non-allergic, hence proving its advantage of not producing any harmful hypersensitivity reaction in the body. It is basic in nature and has low molecular weight hence suitable for any route of administration except oral. Its half-life in *E. coli* is >10 hours, hence can easily be cultured and extracted. It is thermally stable as indicated by instability index. It has various B-cell epitope stimulating site and molecular docking with toll like receptor TLR-7 shows that it binds easily it without any repulsive Van der Waal forces. Toll-like receptor 7 which induces immune response against ss-RNA organisms will elicit an immune response against this sequence considering it be an active virus and thus fulfilling its purpose as a vaccine. 89.2% of the residues were in the favoured region of Ramachandran plot. These points make it favourable for in-vitro trials and further refinement. All these studies were on web-tool prediction servers designed for such type of studies. Because of the high similarity of hepatitis C genome to hepatitis G genome, it is highly probable that this peptide sequence might act as both hepatitis C and hepatitis G vaccine [8]. Patients with past or current HGV

infection have higher CD4+ lymphocyte counts and better AIDS-free survival rates [9][10][11]. This peptide sequence might cause a breakthrough in the treatment of HIV without exposing them to develop hepatitis. However, since they work on growing databases, these cannot give a complete surety for success in future stages. Along with the advantage of the study, there are some limitations. 22% of the predicted secondary structure is disordered. Instead of 98% proteins being in the favourable region of Ramachandran plot only 89.2% of them is present. Advanced molecular dynamic simulations were not performed like RMSD (root mean square deviation). These disadvantages need to be overcome with better resources but these early results do serve as a guiding path to build future work upon it.

Conclusion

This study has highlighted a potential candidate fulfilling its purpose as hepatitis C vaccine, hepatitis G vaccine and HIV management addition. In-depth studies and refinement might serve to be successful since it has very good results at such an early stage. It is needed to be validated experimentally.

Declarations

Conflict of interest:

The author declares no conflict of interest.

Source of funding:

Nil.

Ethical consideration

Not required.

References

1. Rappuoli R. Reverse vaccinology. *Current opinion in microbiology*. 2000 Oct 1;3(5):445-50.
2. Ali A, Khan A, Kaushik AC, Wang Y, Ali SS, Junaid M, Saleem S, Cho WC, Mao X, Wei DQ. Immunoinformatic and systems biology approaches to predict and validate peptide vaccines against Epstein–Barr virus (EBV). *Scientific reports*. 2019 Jan 24;9(1):1-2.
3. Khan S, Khan A, Rehman AU, Ahmad I, Ullah S, Khan AA, Ali SS, Afridi SG, Wei DQ. Immunoinformatics and structural vaccinology driven prediction of multi-epitope vaccine against Mayaro virus and validation through in-silico expression. *Infection, Genetics and Evolution*. 2019 Sep 1; 73:390-400.

4. Sharp K, Dange S. Application of In-Silico Reverse Vaccinology for Designing Multi-Epitope Vaccine Against Coronavirus (preprint available on ChemRxiv) <https://doi.org/10.26434/chemrxiv.12345653>
5. Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. GenBank. Nucleic acids research. 2011 Jan;39(Database issue):
6. Bank PD. Protein data bank. Nature New Biol. 1971; 233:223.
7. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. UCSF Chimera—a visualization system for exploratory research and analysis. Journal of computational chemistry. 2004 Oct;25(13):1605-12.
8. Kim JP, Fry KE. Molecular characterization of the hepatitis G virus. J Viral Hepat 1997; 4: 77-79
9. Toyoda H, Fukuda Y, Hayakawa T, Takamatsu J, Saito H. Effect of GB virus C/hepatitis G virus coinfection on the course of HIV infection in haemophilia patients in Japan. J Acquir Immune Defic Syndr Hum Retrovirol. 1998; 17:209-13.
10. Sabin CA, Devereux H, Kinson Z, Griffioen A, Brown D, Dusheiko G, et al. Effect of coinfection with hepatitis G virus on HIV disease progression in hemophilic men. J Acquir Immune Defic Syndr Hum Retrovirol. 1998;19: 546-8.
11. Lefre` re JJ, Roudot-Thoraval F, Morand-Joubert L, Petit JC, Lerable J, Thauvin M, et al. Carriage of GB virus C/hepatitis G virus RNA is associated with a slower immunologic, virologic, and clinical progression of human immunodeficiency virus disease in coinfecting persons. J Infect Dis. 1999;179: 783-9.

Figures

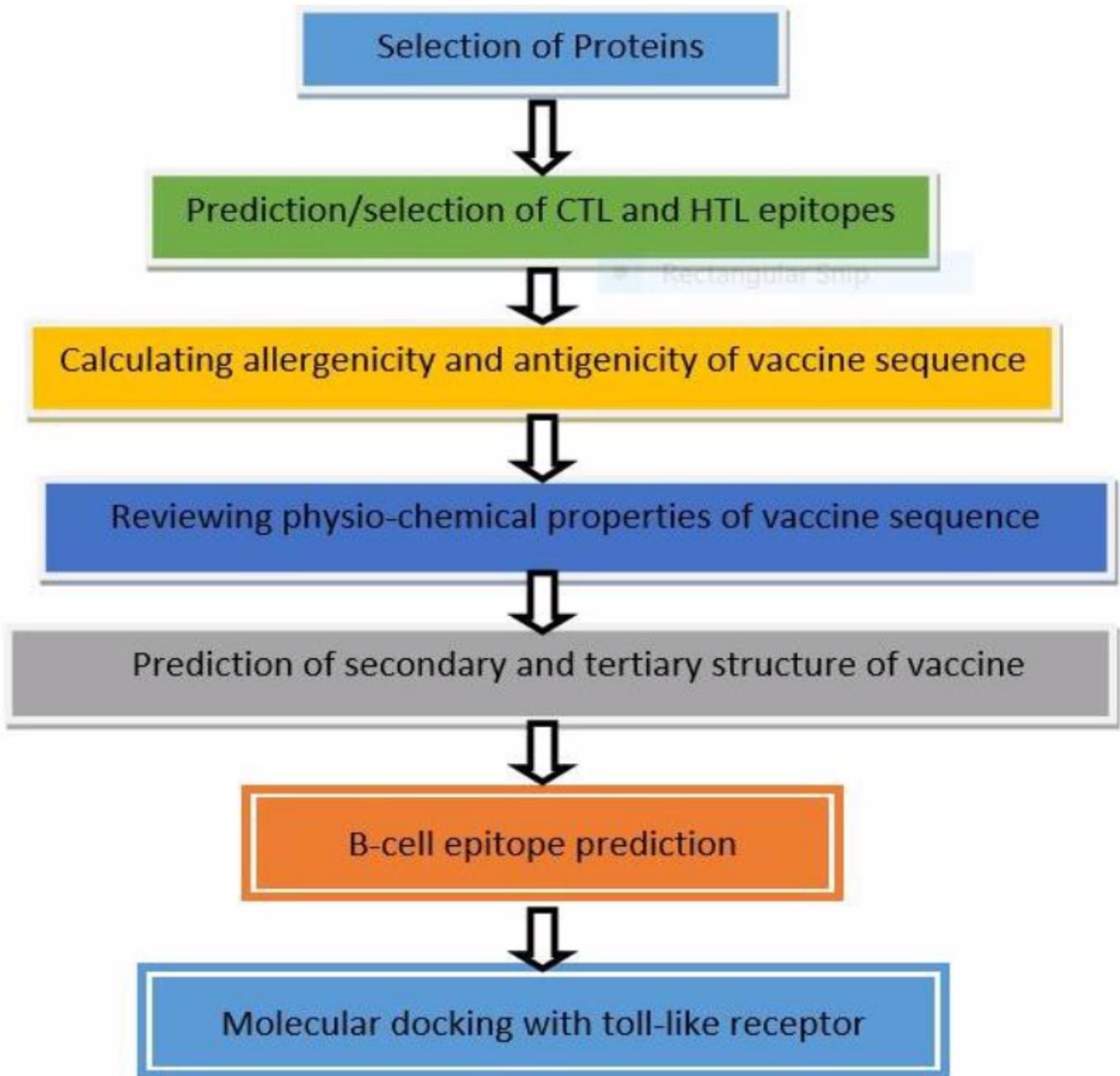


Figure 1

Methodology summary from a similar previous study. Reproduced with permission from the authors.



Figure 2

Secondary structure of vaccine sequence.

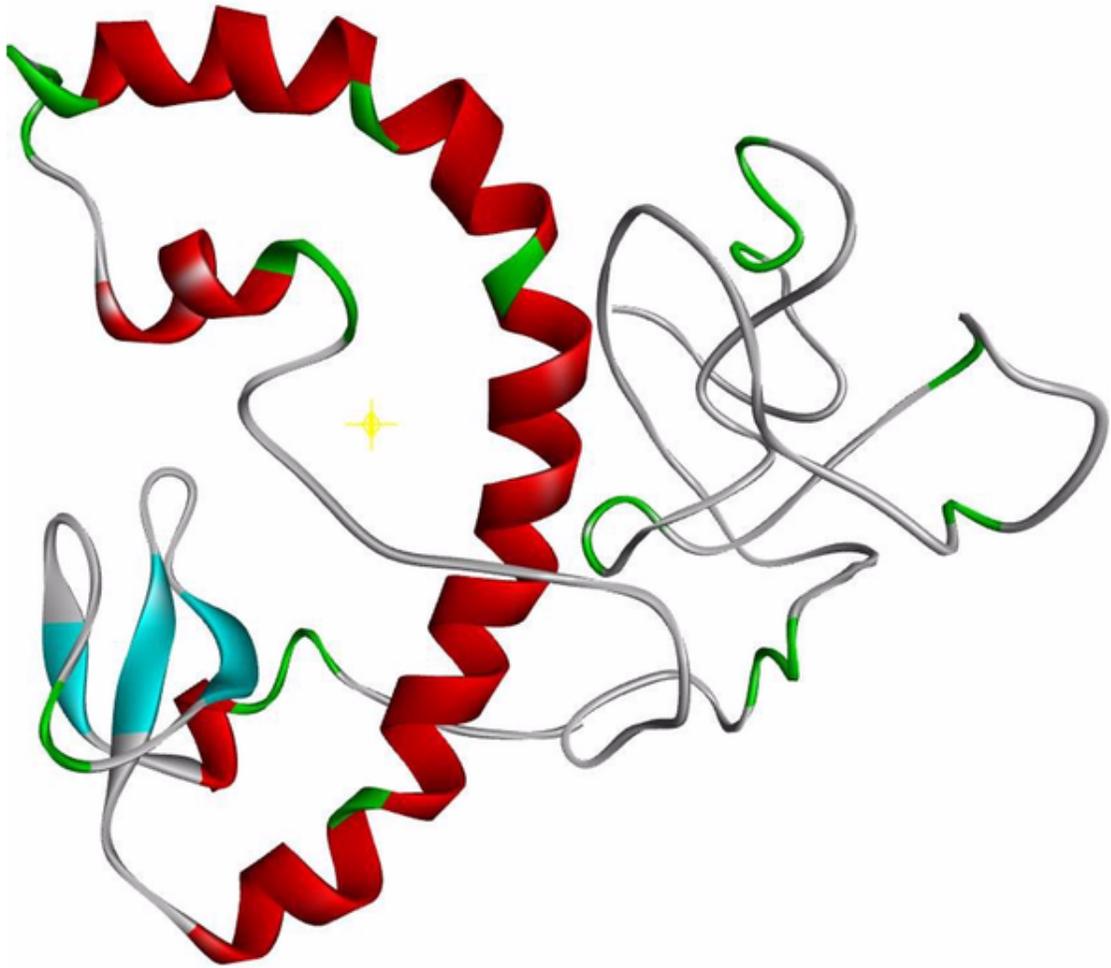


Figure 3

Model 4 tertiary structure.

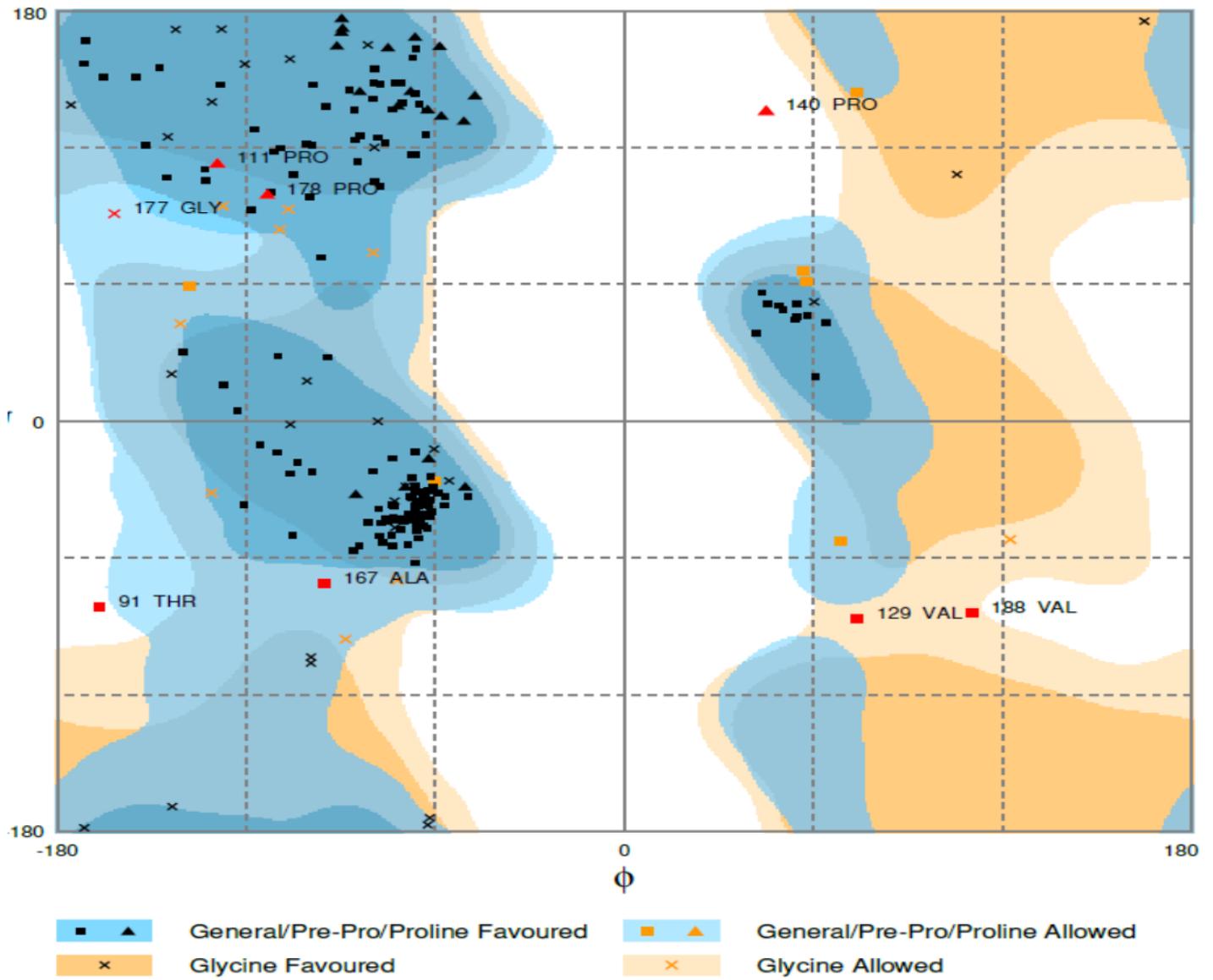


Figure 4

Ramachandran Plot analysis of the final vaccine tertiary structure.

| No. | Chain | Start | End | Peptide | Number of residues | Score |
|-----|-------|-------|-----|-----------------------------|--------------------|-------|
| 1 | - | 75 | 99 | TDAAVAIIQAAYDVALETELYGPGPG | 25 | 0.815 |
| 2 | - | 40 | 51 | CCRRKEAAAKAV | 12 | 0.782 |
| 3 | - | 159 | 175 | GAVFFSGLAPLRMHPDG | 17 | 0.692 |
| 4 | - | 1 | 28 | GIINTLQKYICRVRGGRCVLSCLPKEE | 28 | 0.662 |
| 5 | - | 194 | 214 | TGPGPGVFFSGLAPLRMHPDV | 21 | 0.628 |
| 6 | - | 138 | 145 | PGPLYQAG | 8 | 0.606 |

Figure 5

B-cell linear epitopes predicted by Ellipro suite

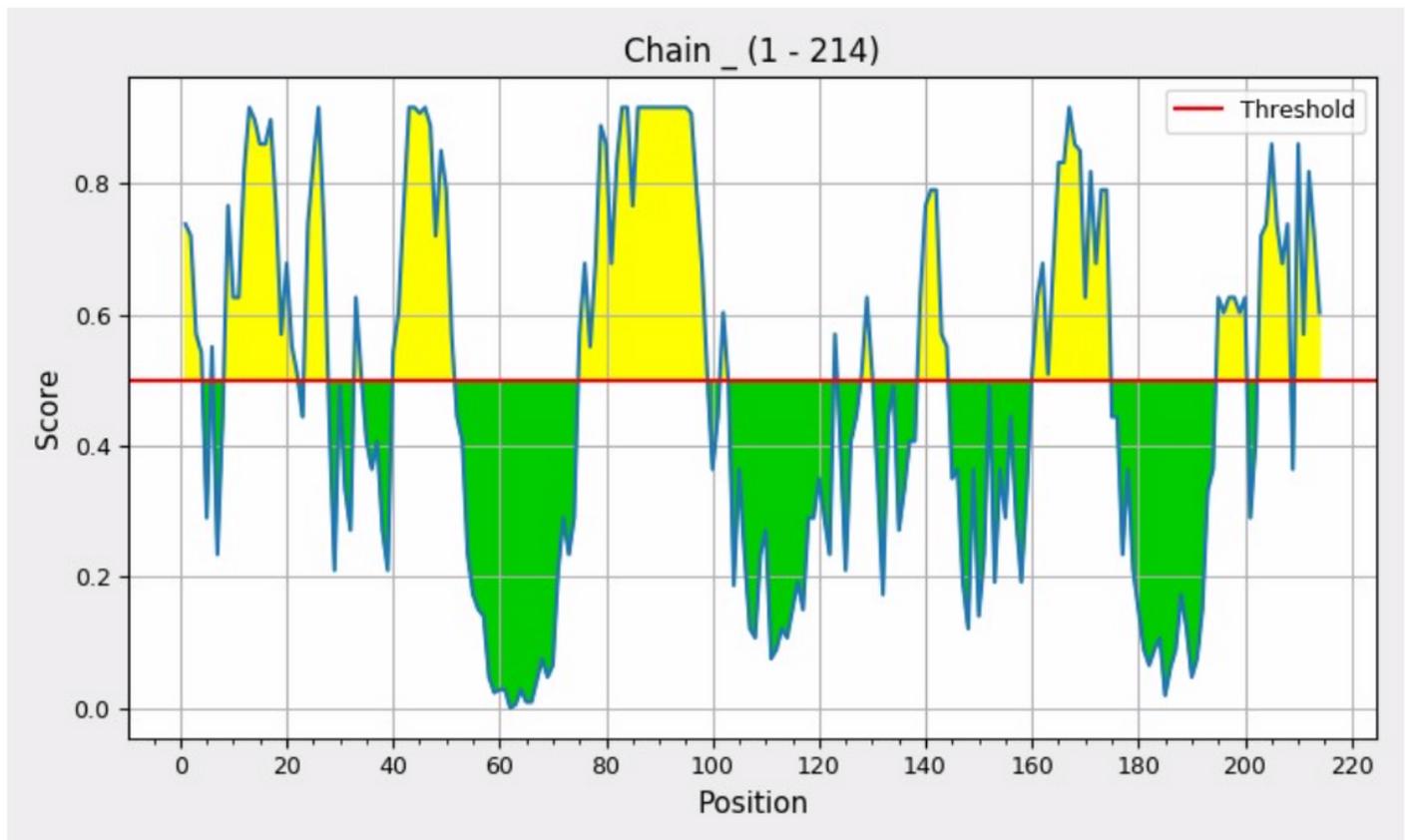


Figure 6

Epitope score chart from Ellipro suite

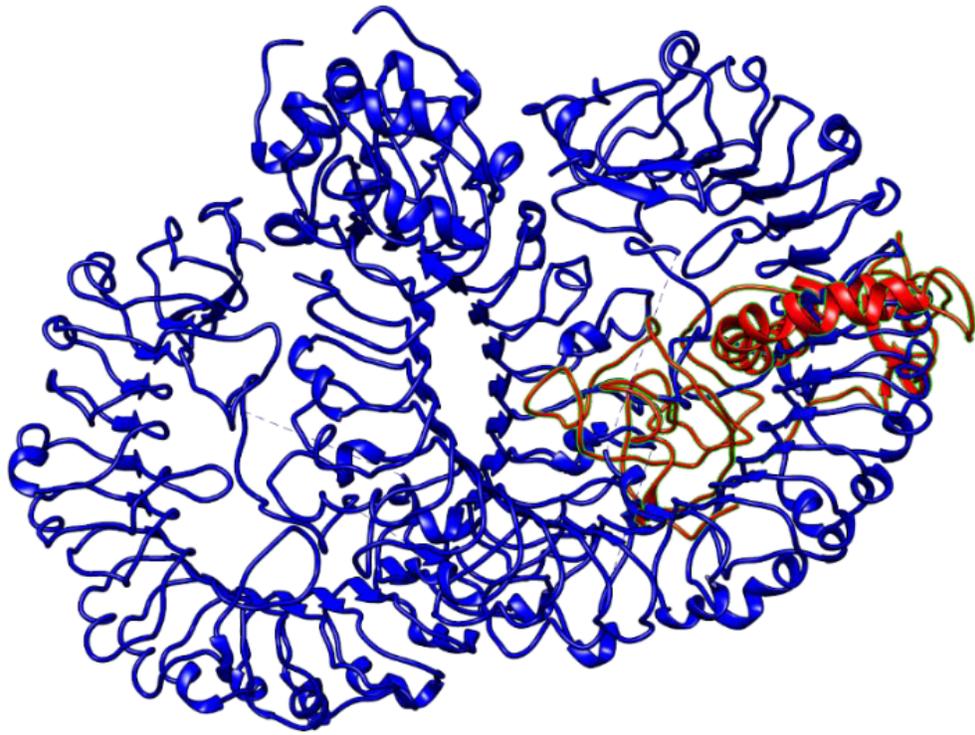


Figure 7

Peptide model(red) docked with Toll-like receptor 7(blue).