ISSN: 0975-3583, 0976-2833 VOL 12, ISSUE 03, 2021

IN –SILICO DESIGN OF EPITOPE BASED PEPTIDE VACCINE AGAINST BARMAH FOREST VIRUS THROUGH IMMUNE-INFORMATIC APPROACH

[1]Melike Sravani,[2]M.Lakshmi Durga,[3]B.Suranjan,Bala Yaswanth Kumar.S,[4]A.Deekshit,[5] Kanaka Durga Devi.N*
KVSR Siddhartha College of Pharmaceutical Sciences, Vijayawada, Andhra Pradesh, India.
*Corresponding author:
Dr. Kanaka Durga Devi.Nelluri
Assoc.Professor,
KVSR Siddhartha College of Pharmaceutical Sciences,
Vijayawada-520010, Andhra Pradesh, INDIA.
E-mail: nelluriss@gmail.com

Abstract:
BFV mainly causes the poly arthritis, occurs in between the ages of 20 to 60. Currently there is no vaccine or drugs for the treatment. We describe an In-silico multi epitope vaccine targeting BFV membrane glycoproteins and predict B and T cell epitopes for immunogenicity, allergenicity, toxicity and conservancy. Epitopes were screened by docking with HLA alleles to generate highest immune response. The 3D structure of the vaccine was determined. Results need to be supported in – vitro and in- vivo for efficacy and safety.
Key -Words: BFV, Multi Epitope Vaccine, Glyco Protein, Immuno Informatics, Vaccine Design.

INTRODUCTION :
Barmah Forest (BF) virus is a mosquito-borne alphavirs, found only in Australia(1), which causes outbreaks of polyarthritis in humans. The disease is very similar to epidemic polyarthritis caused by infection with Ross River virus, another Australian alphavirus. BF virus was first isolated from mosquitoes in the State of Western Australia in 1989. After this, small clusters of human cases were diagnosed in the arid northern and central regions of Western Australia in 1992, and the first substantial outbreak of human disease due to infection with BF virus (BF virus disease) occurred in the southwestern region of the state during the spring and summer (September-March) of 1993-94 (2). The ecology of Australian arboviruses that cause human disease, including BF virus, has recently been reviewed (3). BF virus was first isolated from Culex annulirostris mosquitoes collected at the Barmah Forest in northern Victoria (southeastern Australia) in 1974 (4). It was first shown to infect humans in New South Wales (central-eastern and southeastern Australia) in 1986 (5) and was reported as a cause of clinical disease in humans in 1988 (6). The most common clinical features include polyarthritis, arthralgia, myalgia, fever, rash, and lethargy (7); in some cases, symptoms may persist for more than 6 months.
Currently there is no treatment . prediction of cytotoxic and helper T-lymphocytes
T cell epitopes play a vital role in vaccine design as it significantly reduces the amount of time required for vaccine development. we describe an in-silico multi – epitope for barmah forest virus vaccine targeting virus membrane glycoprotein polyprotein.

MATERIALS AND METHODS :

- **Amino acid sequence retrieval and Secondary structural analysis of target protein :**

  Amino acid sequence of membrane glycoprotein polyprotein of the virus was retrieved from National Centre of target protein was predicted by online server VaxiJen v2.0 (http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html)at threshold 0.4 (Doytchinova and Flower for Biotechnology Information (NCBI) database accession number QIN54116.1in FASTA format. Antigenicity 2007).

- **Prediction of B cell and T cell epitopes**:

B cell lymphocyte epitopes were predicted from target protein sequence by Bepipred linear epitope prediction Immune-Epitope Database and Analysis-Resource (https://www.iedb.org/). .From predicted epitopes, 12 with peptide length > 9-mer were chosen for further analysis.MHC class-I restricted CD8+ CTL epitopes of the target protein sequence were predicted using NetCTL 1.2 server(Larsen et al. 2007). NetCTL 1.2server predicts 9-mer CTL epitopes for the 12 most frequently occurring HLA class I alleles in humans such as A1, A2, A3, A24, A26, B7, B8, B27, B39, B44, B58, and B62. Threshold for proteasomal C-terminal cleavage, Transporter Associated with Antigen Processing (TAP)transport efficiency and epitope identification was set at 0.15, 0.05 and 0.75 respectively

ISSN: 0975-3583, 0976-2833 VOL 12, ISSUE 03, 2021

(Ullahet al., 2020). From the 216 predicted epitopes, 10 epitopes with affinity towards multiple HLAclass 1 alleles were screened for further analysis.

NetMHCIIpan 3.2 server predicted MHC class-II restricted CD4+ HTL epitopes (Jensen et al. 2018). 15-mer HTL epitopes were obtained for HLA Class II DRB1 alleles: 01:01, 03:01, 04:01, 07:01, 08:03, 10:01, 11:01, 12:01, 13:02, 14:01and 15:01 with threshold for strong binder and weak binder at 2 and 10% respectively. These HLA class II alleles were selected as they cover 95% of world's population (Chauhan et al. 2019).Among the 101 strong binding epitopes, 7 epitopes were found to show strong binding towards multiple allelic forms of HLA.

Prediction of epitope properties:

Antigenicity of screened epitopes was predicted by VaxiJen v2.0 (Doytchinova and Flower2007), allergenicity with AllerTop v.2.0

(https://www.ddg-pharmfac.net/AllerTOP/) and toxicity by ToxinPred

(http://crdd.osdd.net/raghava/toxinpred/) (Gupta et al. 2013). TMHMM v.2.0server was utilised to determine the transmembrane topology (http://www.cbs.dtu.dk/services/TMHMM/).. Epitope possible antigens, non-allergic or toxic were used.

Epitope conservancy and population coverage:

Epitope linear sequence conservancy of the predicted B and T cell epitopes was performed using IEDB conservation-analysis-tool (http://tools.iedb.org/conservancy/) to check variability and degree of conservancy in protein sequences from various countries. The epitopes with100% conservancy were analysed further. For the population coverage analysis, sequences of (http://tools.iedb.org/population/)predicted CTL epitopes along with their restricted MHC alleles were submitted to IEDB population coverage analysis tool maintaining the other default parameters.

3D structure modelling and molecular docking:

The de novo 3 dimensional structures of the selected T cell epitope sequences were modeled using an online PEPFOLD 3 server at RPBS MOBYL portal (Maupetit et al. 2009). 10 models of each peptide sequence were retrieved in PDB format. The X-Ray Diffraction structures with PDBID 4U6Y and 1BX2 belonging to HLA-A*01:01 (HLA class I allele) and HLA-DRB1* 15:01 (HLAclass II allele) were retrieved from Protein data bank (PDB).

Structure modelling, refinement and validation of vaccine:

Secondary structure of final vaccine construct was predicted in SOPMA secondary structure prediction method tool by keeping output width, similarity threshold and window width at 70,8 and 17 respectively(Geourjon and Deléage 1995). 3D structure modelling was done in an online server trRosetta (Yang et al. 2020) (https://yanglab.nankai.edu.cn/trRosetta/).Subsequently, the refinement of the generated 3D model of vaccine was carried out in Galaxy Refine. Validation of the refined 3D vaccine model was done in PROCHECK server by analysing Ramachandran Plot.

Toll like receptor-8 and vaccine docking:

A number of studies have stated that immune response against RNA virus is carried out by Toll like receptot-8 (TLR-8) (Lester and Li 2014). So, 3D structure of vaccine construct was docked against TLR-8. The X-Ray Diffraction structure of TLR-8 (PDB ID: 3W3G) with resolution 2.3Awas retrieved from Protein Data Bank (PDB). Docking analysis was performed in HawkDock server (http://cadd.zju.edu.cn/hawkdock/) to ensure that the constructed vaccine generates immune response.

Molecular dynamic simulation of docked complex:

Molecular dynamic simulation of receptor vaccine complex was executed in anonline server iMODS (http://imods.chaconlab.org/).. Molecular dynamic simulations were performed in order to ensure the stability and analyze the physical motion of atoms and macromolecules in interaction of receptor-vaccine complex.

RESULTS :

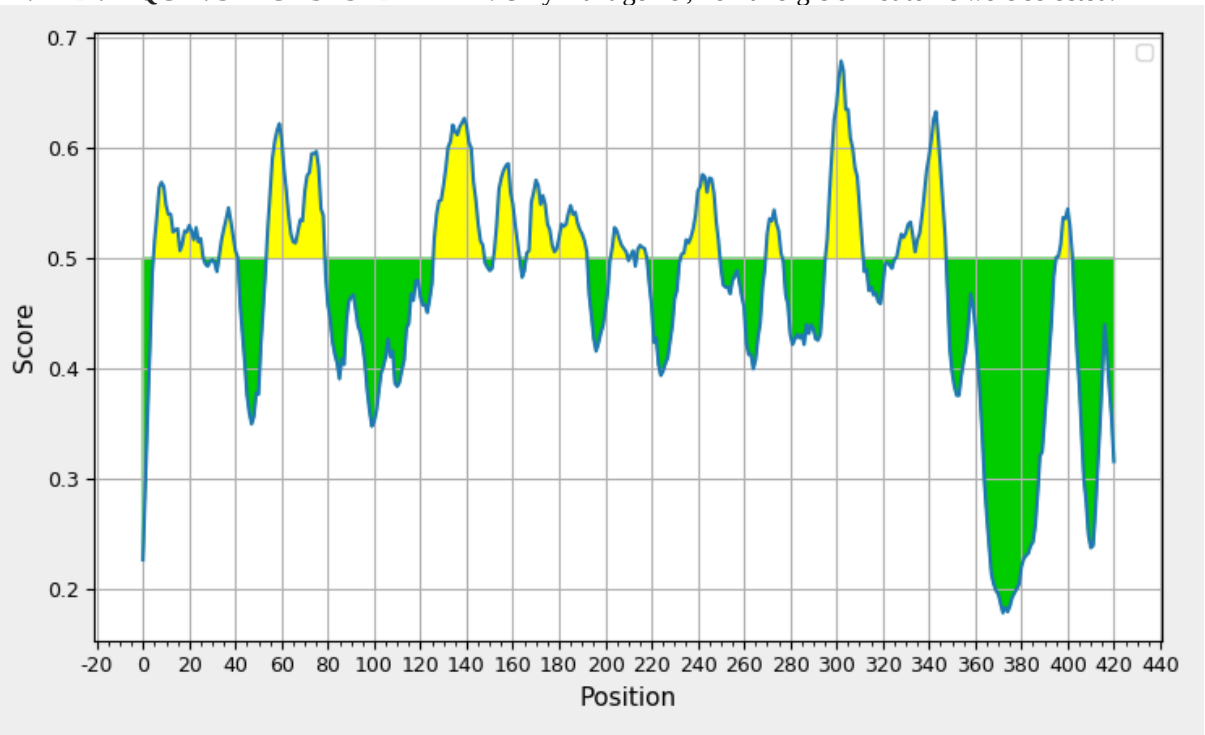
Structural analysis of target proteins:

ISSN: 0975-3583, 0976-2833 VOL 12, ISSUE 03, 2021

Membrane glycoprotein polyprotein sequence of barmah forest virus from NCBI comprised two long open reading frames in the RNA encode a non structural polyprotein of **2411 amino acids** and a structural polyprotein of 1239 amino acids, respectively. The BFV envelope protein E2 is unique among sequenced alphaviruses in having no N-linked glycosylation sites; E1 carries two glycosylation sites. Antigenicity score of target protein was**0.5954** .**Molecular weight:** 46097.73, no of positively charged residues are 39, no of negatively charged residues are 33,**Aliphatic index:** 78.98, **Grand average of hydropathicity (GRAVY):** -0.170.

B cell epitopes prediction:

We generated 14 epitopes of which 9 were screened by peptide length >9 and conservancy of100% present in viruses sequenced in various countries (Ullah et al., 2020). Highest antigenicity was shown by **HVPPDVPIQGLVSNTGKSYSLDPKTKT**. Only 2 antigenic , non-allergic or not toxic were selected.



B Cell epitope prediction

ISSN: 0975-3583, 0976-2833 VOL 12, ISSUE 03, 2021

Table 2 : B cell antigenicity, allergenicity, toxicity, topology

S.NO	PEPTIDE SEQUENCE	antigenicity	allergenicity	toxicity	topology
1.	EHVWSDADD	**NON-ANTIGEN -0.3287**	allergen	Non-Toxin	outside
2.	YSHQTDLTREE	**NON-ANTIGEN 0.0783**	**NON-ALLERGEN**	Non-Toxin	inside
3.	SLGKDPNHSQEWIDTP	**NON-ANTIGEN 0.0783**	**ALLERGEN**	Non-Toxin	outside
4.	WQYNSQYVPRSEVTEVK	**ANTIGEN 0.7558**	**ALLERGEN**	Non-Toxin	inside
5.	EAYKATRPYIGWCADCGLAGSC	**NON-ANTIGEN 0.2948**	**NON-ALLERGEN**	Toxin	inside
6.	**SHKPRFIGNEKSPAPTGHKTRI**	**ANTIGEN 0.5304**	**NON-ALLERGN**	**Non-Toxin**	**outside**
7.	VWGNNNPVRLWAQKSSSSSAHG	**NON-ANTIGEN -0.3671**	**ALLERGEN**	Non-Toxin	outside
8.	IAKSNTINHAKIRYMGANGVQEAER	**ANTIGEN 0.4913**	**ALLERGEN**	Non-Toxin	inside
9.	**HVPPDVPIQGLVSNTGKSYSLDPKTKT**	**ANTIGEN 0.5791**	**NON-ALLERGN**	**Non-Toxin**	**outside**

ISSN: 0975-3583, 0976-2833 VOL 12, ISSUE 03, 2021

Table 3 :HLA CLASS 1 alleles and super types , antigenicity, toxicity, toxicity and topology

.

POSITION	HLA CLASS 1 ALLELES & SUPERTYPES	PEPTIDE SEQUENCE	ANTIGENICITY	allergenicity	toxicity	topology
232	A1, A3, A26, B62,	TVWQYNSQY	**NON-ANTIGEN 0.3975**	**NON-ALLERGEN**	Non-Toxin	inside
398	B27,B39, B62	YQLAPGAQL	**ANTIGEN 0.8529**	**ALLERGEN**	Non-Toxin	outside
277	B58, B62	RLGEVEFHF	**ANTIGEN 2.3136**	**ALLERGEN**	Non-Toxin	outside
374	A2,B8,B62	LLIVISSGL	**NON-ANTIGEN 0.3247**	**ALLERGEN**	Non-Toxin	outside
361	**B7, B8,B39**	**YPYWTITVL**	**ANTIGEN 1.3457**	**NON-ALLERGEN**	**Non-Toxin**	**outside**
60	**A1,A3,A26**	**TINHAKIRY**	**ANTIGEN 0.6500**	**NON-ALLERGEN**	**Non-Toxin**	**inside**
353	A1,B58	IASHYYDLY	**ANTIGEN 0.4534**	**ALLERGEN**	Non-Toxin	outside
318	A1,B62	QVGAEGVEY	**ANTIGEN 1.1633**	**ALLERGEN**	Non-Toxin	outside
218	**B44, B62**	**VEFHFHPMY**	**ANTIGEN 2.1256**	**NON-ALLERGEN**	**Non-Toxin**	**outside**
350	**A1, B62**	**PISIASHYY**	**ANTIGEN 0.8902**	**NON-ALLERGEN**	**Non-Toxin**	**inside**

ISSN: 0975-3583, 0976-2833 VOL 12, ISSUE 03, 2021

Table 4 : HLA CLASS 2 alleles antigenicity, allergenicity, toxicity and topology.

POSITION	PEPTIDE SEQUENCE	SUB TYPES	antigenticity	allergenicity	toxicology	topology
63	HAKIRYMGANGVQA	DRB1_0101 DRB1_0401 DRB1_0701 DRB1_0803 DRB1_1001 DRB1_1201 DRB1_1302 DRB1_1401 DRB1_1454	**ANTIGEN 0.6815**	**ALLERGEN**	Non-Toxin	inside
62	NHAKIRYMGANGVE	DRB1_0101 DRB1_0401 DRB1_0701 DRB1_0803 DRB1_1001 DRB1_1201 DRB1_1302 DRB1_1401 DRB1_1454	**ANTIGEN 0.5174**	**ALLERGEN**	Non-Toxin	inside
64	**AKIRYMGANGVQEAE**	**DRB1_0101 DRB1_0701 DRB1_0803 DRB1_1001 DRB1_1201 DRB1_1302 DRB1_1401 DRB1_1454**	**ANTIGEN 0.7000**	**NON-ALLERGEN**	**Non-Toxin**	**inside**
61	INHAKIRYMGANGQ	DRB1_0101 DRB1_0701 DRB1_1001 DRB1_1201 DRB1_1302 DRB1_1401 DRB1_1454	**ANTIGEN 0.6469**	**ALLERGEN**	Non-Toxin	inside
128	HKPRFIGNEKSPAT	DRB1_0803 DRB1_1101 DRB1_1302	**NON-ANTIGEN 0.3006**	**ALLERGEN**	Non-Toxin	inside
170	PDVPIQGLVSNTGS	DRB1_0101 DRB1_0401 DRB1_1001	**NON-ANTIGEN -0.0658**	**ALLERGEN**	Non-Toxin	outside

Population coverage by selected MHC class-I epitope

ISSN: 0975-3583, 0976-2833 VOL 12, ISSUE 03, 2021

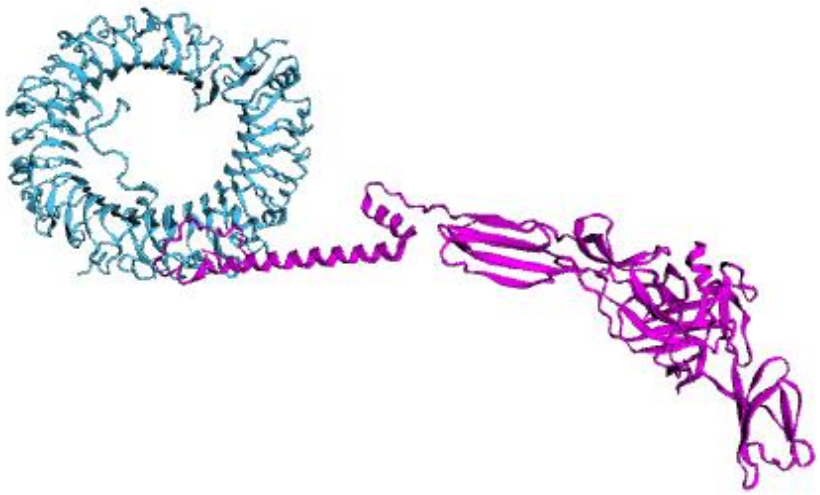
population coverage

POPULATION	COVERAGE
World	88.78 %
East Asia	87.9%
North East Asia	57.77%
South Asia	68.09%
South East Asia	71.06%
South West Asia	70.15%
Europe	96.11%
East Africa	58.06%
West Africa	59.5%
Central Africa	52.00%
North Africa	67.52%
South Africa	58.35%
North America	88.03%
South America	63.02%
Oceania	79.51%

Protein-protein docking and molecular dynamics:

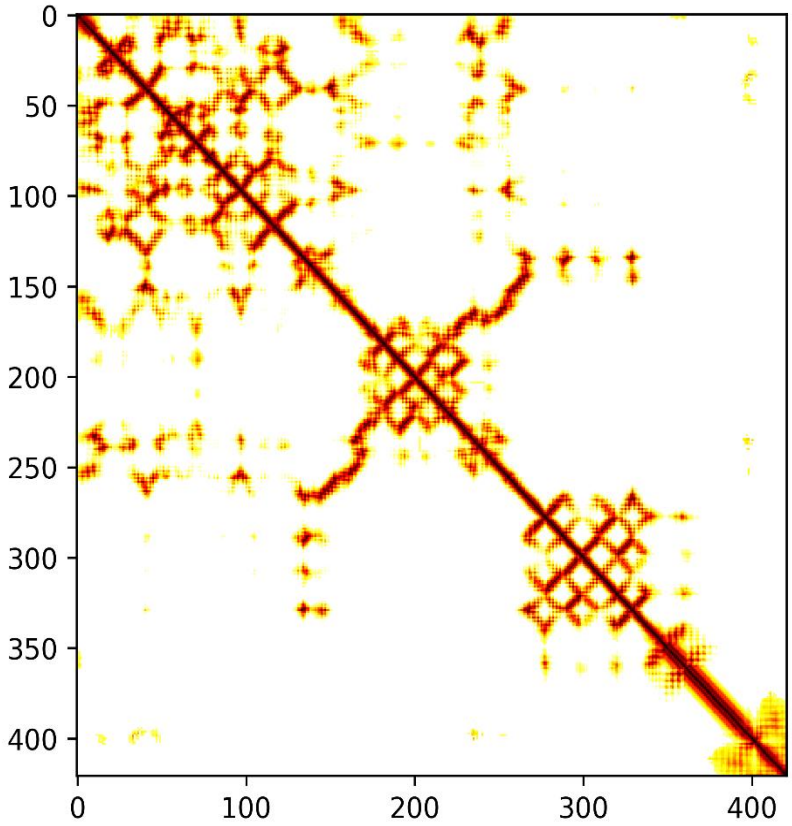
ISSN: 0975-3583, 0976-2833 VOL 12, ISSUE 03, 2021

The vaccine was docked with TLR-8 to generate immune responses with 10 models by HawkDock server .



ISSN: 0975-3583, 0976-2833 VOL 12, ISSUE 03, 2021

TLR-8 vaccine interaction predicted by HawkDock server. The vaccine is represented in pink and blue color.



Ramachandran plot .

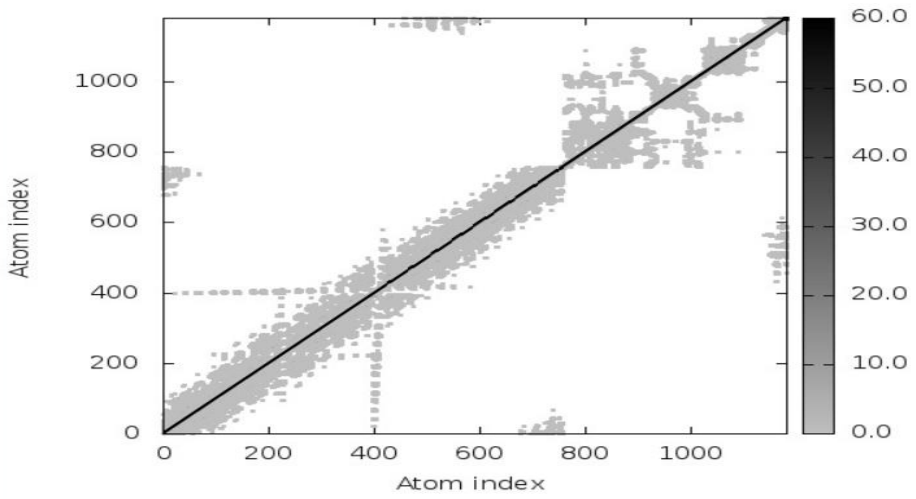
Later, molecular dynamic simulation studies were carried out on TLR-8 and vaccine docked in iMODS server. Elastic network by Figure 13 A below. Figure 13 B shows B-factor graph comparing PDB and NMA (normal mode analysis) of the receptor-ligand docked complex.

Eigenvalues represent energy required to deform the structure: we found TLR-8 and vaccine docked complex Eigenvalue of 3.874705e-06 (Figure 13 C). Variance of each mode is inversely proportional to eigenvalue (Figure 13 D) and covariance map designates coupling between pairs of residues.

ISSN: 0975-3583, 0976-2833 VOL 12, ISSUE 03, 2021

Elastic network

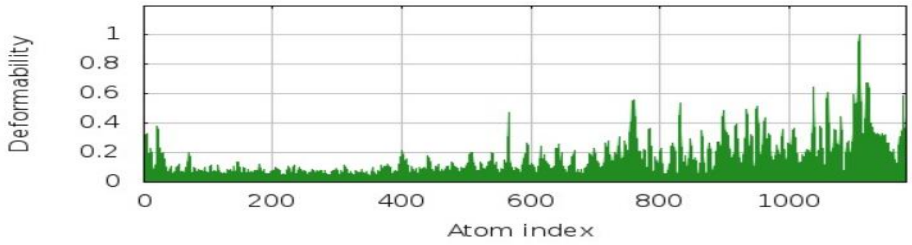
The elastic network model defines which pairs of atoms are connected by springs. Each dot in the graph represents one spring between the corresponding pair of atoms. Dots are colored according to their stiffness, the darker grays indicate stiffer springs and vice versa.



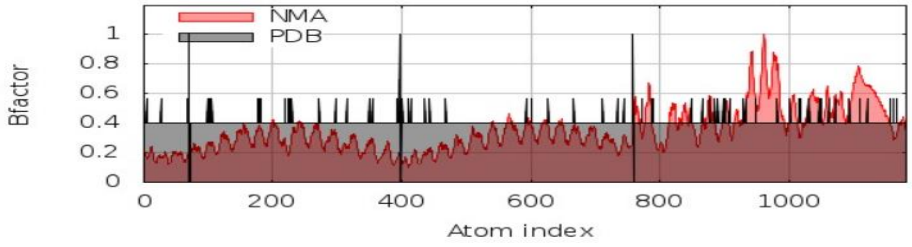
Elastic network

B-factor/Mobility

The main-chain deformability is a measure of the capability of a given molecule to deform at each of its residues. The location of the chain 'hinges' can be derived from high deformability regions.



The experimental B-factor is taken from the corresponding PDB field and the calculated from NMA is obtained by multiplying the NMA mobility by (8pi^2□). Be aware that many PDB files of averaged NMR models contain no B-factors (actually, the B-factor column gives an averaged RMS).



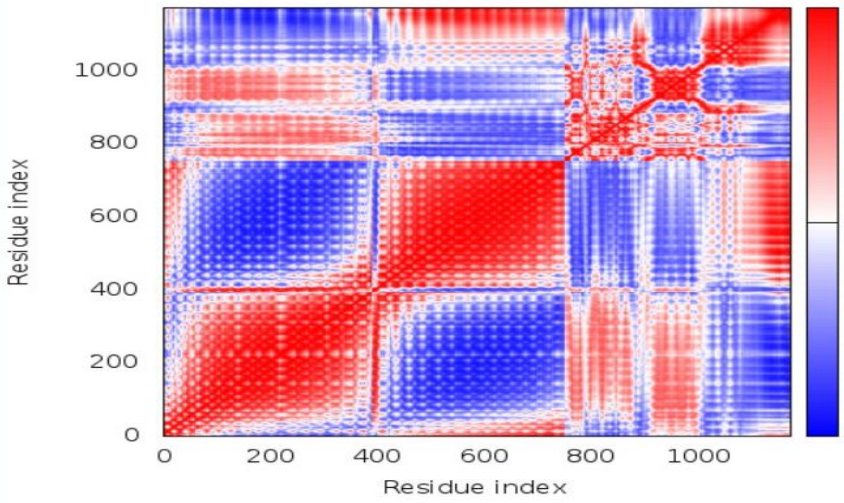
B – factor / Mobitity

Eigen Values and Variance Eigen Values and Variance

ISSN: 0975-3583, 0976-2833 VOL 12, ISSUE 03, 2021

Covariance map

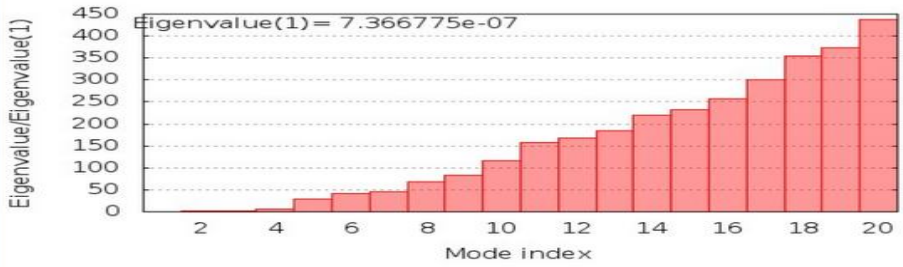
Covariance matrix indicates coupling between pairs of residues, i.e. whether they experience correlated (red), uncorrelated (white) or anti-correlated (blue) motions.



The correlation matrix is computed using the Cα Cartesian coordinates and the equation 2 in: Ichiye T, Karplus M. Collective Motions in Proteins: a Covariance Analysis of Atomic Fluctuations in Molecular-Dynamics and Normal Mode Simulations. Proteins 11 205-217 (1991).

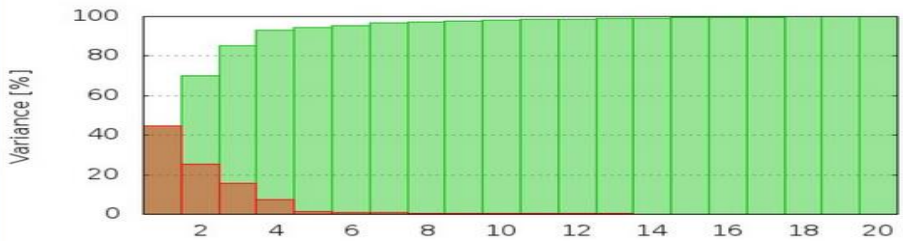
Eigenvalues

The eigenvalue associated to each normal mode represents the motion stiffness. Its value is directly related to the energy required to deform the structure. The lower the eigenvalue, the easier the deformation.



Variance

The variance associated to each normal mode is inversely related to the eigenvalue. Colored bars show the individual (red) and cummulative (green) variances.



DISCUSSION :

The National Institute of Allergy and infectious diseases classifies thrombocytopenia syndrome virus as a category C emerging pathogen with potential for high morbidity and mortality. https://www.niaid.nih.gov/research/emerging-

ISSN: 0975-3583, 0976-2833 VOL 12, ISSUE 03, 2021

infectious-diseases-pathogens. Favipiravir was shown to effectively reduce neurological signs if given within 5 days of infection but there are no specific approved antivirals or vaccines (Kato et al., 2016). Kwak et al., (2019) and Kang et al., (2020) developed DNA vaccines and studied their immunogenicity in various animal models but are yet to undergo clinical trials in humans to show their efficacy prior to their approval. Unlike the single epitope based vaccines, using multiple CTL, HTL and B cell epitopes in a vaccine can induce activation of both humoral and cellular immune responses with minimum adverse effects (Chauhan et al., 2019). In our study we designed a multi-epitope based vaccine using an immune informatic approach for BFV from its membrane glycoprotein polyprotein. A number of BCL and T cell epitopes were predicted. Among these, only few epitopes which were antigenic were chosen carefully by screening through numerous immune filters in a sequential manner.

CONCLUSION :

Computer modelling methods aid in large scale screening of peptides with all the possible HLA alleles to procure finest peptides in extensive population. These methods are efficient in decreasing time and money invested in discovering the epitopes with high specificity for vaccine design. The vaccine designed in this study was based on membrane glycoprotein polyprotein of SFTSV through a reverse vaccinology approach. Further experimental validation is necessary to evaluate the therapeutic efficacy and immunogenicity of the vaccine.

ACKNOWLEDGEMENTS:

The authors would like to acknowledge the management and the Principal, KVSR Siddhartha College of Pharmaceutical Sciences for supporting this work.

REFERENCE :

1. Lindsay MD, Smith DW, Johansen C, Mackenzie JS. Barmah Forest virus disease in Western Australia. Western Australian Notifiable Diseases Bulletin 1994; 4: 1-4.
2. Lindsay MD, Johansen C, Smith DW, Wallace MJ, Mackenzie JS. An outbreak of Barmah Forest virus disease in the south-west of Western Australia. Med J Aust. In press.
3. Mackenzie JS, Lindsay MD, Coelen RJ, Hall RA, Broom AK, Smith DW. Arboviruses causing human disease in the Australian zoogeographic region. Arch Virol 1994;136:447-7.
4. Marshall ID, Woodroofe GM, Hirsch S. Viruses recovered from mosquitoes and wildlife serum collected in the Murray Valley of south-eastern Australia, February 1974, during an outbreak of encephalitis. Aust J Exp Biol Med Sci 1982;60:457-70.
5. Vale TG, Carter IW, McPhie KA, James GS, Cloonan MJ. Human arbovirus infections along the south coast of New South Wales. Aust J Exp Biol Med Sci
1986;64:307-9.
6. Boughton CR, Hawkes RA, Naim HM. Illness caused by a Barmah Forest-like virus in New South Wales. Med J Aust 1988;148:146-7.
7. Phillips DA, Murray JR, Aaskov JG, Weimers MA. Clinical and subclinical Barmah Forest virus infection in Queensland. Med J Aust 1990;152:463-6.