

3D Structure Development of Chikungunya Virus and its Structural Characterization – an *In silico* approach

Anuraj Nayariseri^a, Mukesh Yadav^a, D Muralidhara Rao^b, Sonam Kukreja^c, Sheaza Ahmed^c, Pooja Singh^d,

a) Bioinformatics Research Laboratory, Eminent Biosciences, Vijaynagar, Indore – MP India.

b) Dept. Of Biotechnology, Bioinformatics Laboratory Sri Krishnadevaraya University, Anantapur-AP India.

c) Dept of Bioinformatics, Softvision College, Vijaynagar, Indore – Madhyapradesh, India.

d) Dept. of Microbiology, Maharaja Ranjit Singh College, Indore – Madhyapradesh, India.

Correspondance Author: - E-mail: anuraj@eminentbio.com Tel: +91 73142 56149,

Abstract

The 3 Dimensional structure of a protein is an expedient for structure based drug design and identifying the conformational epitopes that are foremost for designing the vaccines. Chikungunya (CHIKV) is a heat sensitive RNA virus which causes a viral disease which is transmitted in humans by human-mosquito-human transmission. Disease mostly found in the tropic and subtropical countries mainly in South India and in few other Asian countries. The primary transmission agents of Chikungunya Virus are yellow fever mosquito (*Aedes aegypti*) and forest day mosquito (*Aedes albopictus*). The forest day mosquito bites during day time and hence day time mosquito bite is the main reason for transmission. The main distinctive of the disease is a fever along with an arthritic type of pain in the joints. 3D Structure of Chikungunya virus is not encountered in the PDB Databank. A Comparative modeling method was used for the prediction of the 3D structure of Chikungunya virus. Modeled structure prvided substantial parameters under Ramachandran plot and stereochemical aspects of main chain and side chains through PROCHECK analysis. Finally the energy minimization protocols yielded a score of - 12063.947 KJ/Mol and RMSD Value of 0.29°A. Structure Validation procedures proffer the use of present modeled structures for structure based drug design and other confirmative applications.

Keywords- Chikungunya, Virus, Comparative modeling, Ramachandran Plot, Procheck Analysis, RMSD.

Background

Chikungunya is an alphavirus, which carried by the *Aedes aegypti* and spreaded through stagnated water [1]. Due to the heavy monsoon and floods, the modern epidemic of chikengunia is found in south Indian states like Kerala, Tamilnadu, Andhrapradesh and Karnataka [2]. The main symptoms include severe temperature, body pain, pain in all the major joints in legs and hands with swelling, due to arthritis affecting multiple joints. If Forest day mosquito bites an infected person, Then the virus gets consign to its body. The virus does not infect the mosquito itself, but on the adverse, the mosquito bites a healthy person, the enhanced virus is now transferred into the body of the healthy person. Thus the healthy person attains affected [3]. The Chikungunya virus depends upon the process of viral tropism for its action. Once the virus enters into a host it searches for susceptible cells. As soon as it finds one, the virus fuses with the cell membrane of the cell and it proceeds towards the nucleus of the cell. It then enters the nucleus and hence captures the biosynthetic machinery of the cell, and directs the synthesis of viral proteins as well as the replication of the viral genome. Hence mature and infectious progeny of the Chikungunya virus are produced inside the cell [4]. Once it produced, these particles leave the cell and escape into the surrounding tissues, carrying out the process in them as well. When, due to this process of viral replication when the virus particles increase in number manifold, they finally escape into the blood and this leads to the establishment of the disease [4]. The mode of action of Chikungunya virus, by which it causes the disease remain to be investigated in detail and its mechanism of action has not yet been fully characterized accept the fact that it causes major histopathological changes in the skeletal muscle tissue, severe inflammation and necrosis of skeletal muscle[5].The cells which get infected by the CHIKV rapidly undergo cell death due to apoptosis. So it can be said that the CHIKV stimulates cell death, by apoptosis, of skeletal cells, connective tissue of joints, epithelial cells of skin and the cells lining the inner walls of the organs of the body(i.e. endothelial cells) [4].

Principle finding

Protein 3D structure is mandatory to predict the function and drug binding studies [6]. 3D structure of Chikungunya virus is not discovered in Protein Data Bank which is a public repository of Protein 3D Structure. Hence the objective of our work is to construct the 3D structure of Chikungunya virus. Since the adversity of X-Ray Crystallography and other *in vitro* methods for predicting the 3Dimensional structure, we used Homology modeling for the prediction of 3D structure of Chikungunya virus. Homology Modeling or Comparative modeling [16, 18] is the prediction of 3D structure with the help of Homologous or highly similar structure[7][20]. As a limiting parameter of homology modeling, the template undertaken for backbone alignment should have identical amino acids with 30 or >30% when sequentially aligned with query protein [8][19] followed by RMSD value <1°A[8]. The quality of the sequence

alignment between target sequence and template sequence is of afflictive importance. Misplaced gaps, insertions or deletions, will attempt residues to be displaced [9][20]. Therefore the quality of the outcome of Comparative modeling is based on the Alignment between Query sequence and Template Sequence[10].

Methodology

Retrieval of target protein sequence of Chikungunya virus was obtained from NCBI (<http://www.ncbi.nlm.nih.gov/protein>) [ref Accession num: ACJ64501]. Length of the protein has been reported as 440 amino acids and Molecular Weight reported as 47480.91 Daltons. The template protein was searched through Blastp algorithm [11] against PDB Database [12]. The high resolution of 3.00⁰ Crystal Structure Of The Homotrimer Of Fusion Glycoprotein E1 From Semliki Forest Virus. (PDB code 1RER)[13] shows 63% Identity with the target sequence. Hence the same can be used as a template for the development of 3D structure of Chikungunya virus. We used two templates for modeling. i.e 1RER_A chain and 1RER_B Chain (Figure : 2). In order to predict the conserved secondary structure profiles, a multiple sequence alignment program DSSP have been brought into use which identified the corresponding position amino acids in the query sequence of CHIKV and templates 1RER_A chain and 1RER_B Chain. This is a confirmatory measure to proceed in comaprative modeling (Figure : 1)[14]. Amino acid sequence alignment of target and template proteins was derived using the Swiss-PdbViewer. Default parameters were applied and the aligned sequences were inspected and quadrate manually to minimize the number of gaps and insertions. Structure prediction was rack up by homology modeling method, using Deep View/Swiss PDB Viewer and Swiss Model server [15].

Chikungunya virus peptides were visualized as straight linear chain of amino acids by using SPDBV software, by accessing Load Raw Sequence to Swissmodel menu in SPDBV 3.7 version. Chikungunya virus sequences and the templates sequence were bind during the alignment correction by selecting Magic Fit menu in the SPDBV. The aligned residues could be seen on the Alignment window in SPDBV. Insertions in the target are proportional to the template were respond as loops with known structure[16]. The estimation of main-chain structures of the loops were modelled by using an efficient and accurate Monte Carlo approach [17] and once the main-chain structure was modelled, the side-chain atoms were Affiliated from the Control Panel in SPDBV. In current study, side-chain torsional angles were determined to equal also it is more close to those in the template structure. This is particularly useful in ignoring the clashes between side chains in the modelled structure [16][18]. Ultimately, the all atoms in the final model were subjected to a short run of energy minimization to improve unfavorable steric interactions and to optimize the stereochemistry[18].The Final Model has been Visualized In Accerlerys DS Viewpro (Figure : 3).

Model validation

We used PROCHECK [20] to calculate the main-chain torsional angles, i.e. the Ramachandran plot [19][20], for our predicted structures. The final model shows Out of 384 amino acids 52 are glycine and proline. Herby Procheck has predicted 332 amino acids from our predicted model. Out of 332 amino acids 293 (88.3%) are in most favoured region, 37 (11.1%) are in additional allowed region and only 2 (0.6%) are generously allowed region. As a most confirmatory aspect of present struncture modeling, amino acids are found absent in disallowed region (Figure: 4).

Results and Discussion

The reason for choosing Chikungunya virus protein was because of its function on viral attachment at the host cell surface and to alleviate the immune response at the host cell. No vaccine or specific antiviral treatment for Chikungunya has been predicted so far. After performing the Energy minimization of SPDBV, The final 3D Structure has given the energy of -12063.947 KJ/Mol and RMSD Value of 0.29°A. The overall G factor calculated for modeled structure came to be -2 (inside) as compared to the reference value -4 [20]. Further validation favoured by bad contacts analyzed per 100 amino acids revealed that 2.1 when compared with reference value 4.2 [20]. The current study can provide valuable information of protein structure and function. It can be used for finding the target protein and ligand for the treatment of the disease and its epidemic nature as well.

Reference

- [1] Vijayakumar K, Anish TS, Sreekala KN, Ramachandran R, Philip RR. Environmental factors of households in five districts of Kerala affected by the epidemic of Chikungunya fever in 2007. *Natl Med J India*. 2010 Mar-Apr;23(2):82-4.
- [2] Kaur P., Manickam Ponniah, Manoj V. Murhekar, Vidya Ramachandran, Ramakrishnan Ramachandran, Hari Kishan Raju, Vanamail Perumal, Akhilesh C. Mishra, and Mohan D. Gupte. Chikungunya Outbreak, South India, 2006 *Emerging Infectious Diseases*. Vol. 14, No. 10, October 2008.1623-1625.
- [3] Mohan, A., Kiran, D., Manohar, I. C. and Kumar, D. P. 2010. Epidemiology, clinical manifestations, and diagnosis of Chikungunya Fever: lessons learned from the re-emerging epidemic. *Indian J Dermatol* 55 (1):54-63.
- [4] Schwartz, O. Albert, ML. (2010) Biology and pathogenesis of Chikungunya virus. *Nature Reviews Microbiology* July, vol 8, 491-500.

[5] Sarah A. Ziegler, Liang Lu, Amelia P. A. Travassos da Rosa, Shu-Yuan Xiao, AND Robert B. Tesh. An Animal Model for Studying the Pathogenesis of Chikungunya Virus Infection. *Am. J. Trop. Med. Hyg.*, 79(1), 2008, pp. 133-139.

[6] Lengauer, T. Zimmer, R. Protein structure prediction methods for drug design. *Brief Bioinform* (2000) 1 (3): 275-288.

[7] John B, Sali A. Comparative protein structure modeling by iterative alignment, model building and model assessment. *Nucleic Acids Res.* 2003 Jul 15;31(14):3982-92.

[8] Ginalski K. Comparative modeling for protein structure prediction. *Curr Opin Struct Biol.* 2006 Apr;16(2):172-7. Epub 2006 Feb 28.

[9] Walsh AW, Langley DR, Colonno RJ, Tenney DJ (2010) Mechanistic Characterization and Molecular Modeling of Hepatitis B Virus Polymerase Resistance to Entecavir. *PLoS ONE* 5(2): e9195.

[10] Apatoff A, Kim E, Kliger Y (2006) Towards Alignment Independent Quantitative Assessment of Homology Detection. *PLoS ONE* 1(1): e113.

[11] Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol.* 1990 Oct 5;215(3):403-10.

[12] Berman, H. M. (January 2008). *Acta Crystallographica Section A: Foundations of Crystallography* A64 (1): 88-95.

[13] Lescar, J., A. Roussel, M. W. Wien, J. Navaza, S. D. Fuller, G. Wengler, and F. A. Rey. 2001. The fusion glycoprotein shell of Semliki Forest virus: an icosahedral assembly primed for fusogenic activation at endosomal pH. *Cell* 105:137-148.

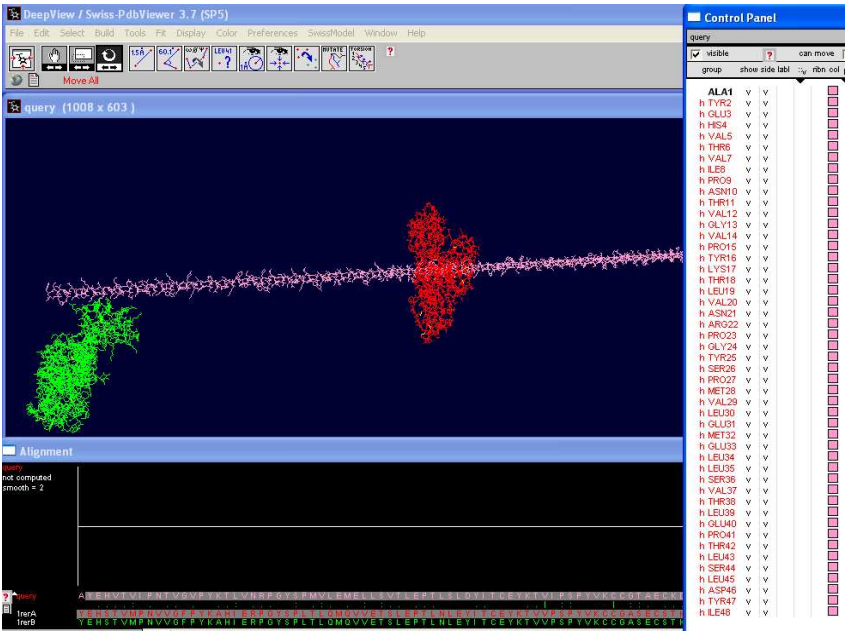
[14] Kabsch, W. and Sander, C. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers.* 1983 22(12):2577-637.

[15] Guex N and Peitsch MC (1997) SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modelling. *Electrophoresis* 18:2714-2723.

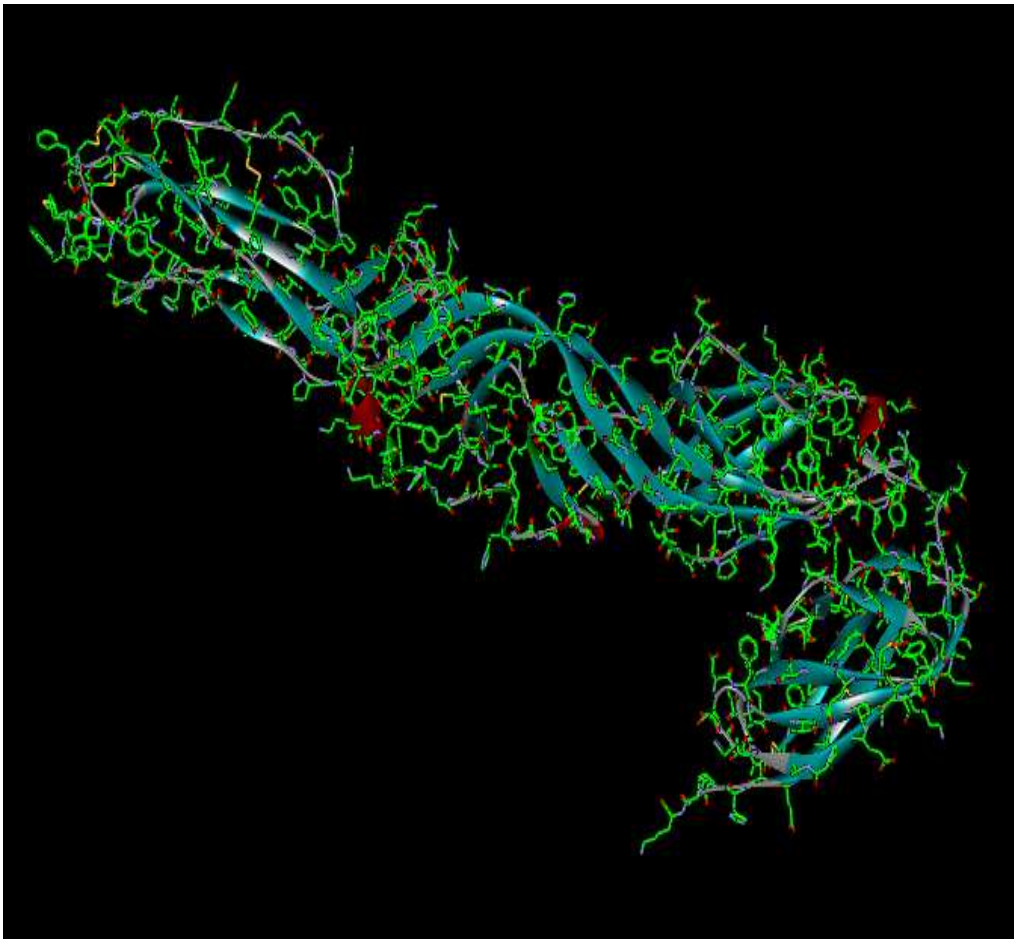
[16] Yadav, M. Anuraj, N. Rajput S. G. Jain, A. Verma, A. Gupta, P. Comparative modeling of 3-oxoacyl-acyl-carrier protein synthase I/II in *Plasmodium falciparum*- A potent target of malaria. *IJBR*, Volume 1, Issue 2, 2009, pp-100-103.

Figure: 01 Shows the Multiple alignment Performed by DSSP (Kabsch and Sander, 1983) with Chickenguniya sequence against two templates which used for homology modeling.

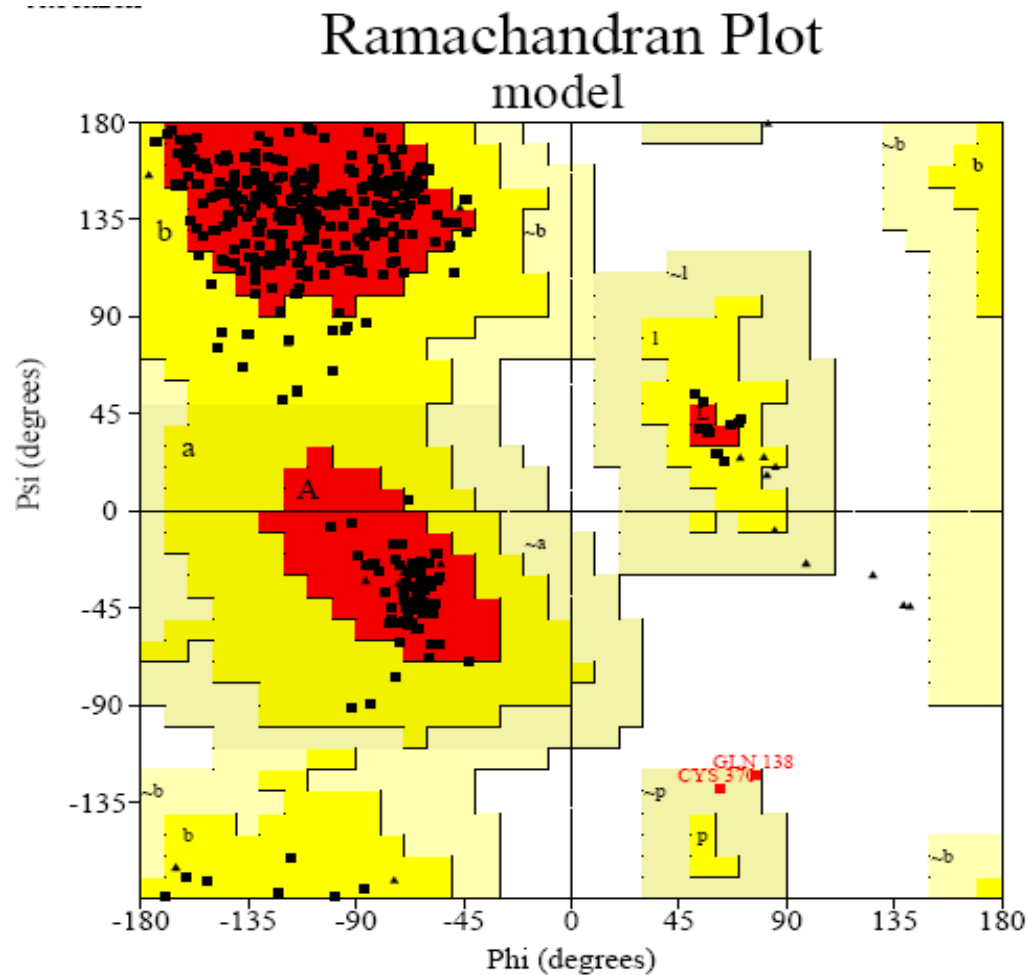
[Figure: 2] Two Templates are Used for Modeling. (1RER_A chain red shows color & 1RER_B Chain green color)



[Figure : 3] The Final Model of Chikungunya virus



[Figure : 4] The Ramachandran Plot produced by Procheck



Residues in most favoured regions [A,B,L]	293	88.3%
Residues in additional allowed regions [a,b,l,p]	37	11.1%
Residues in generously allowed regions [~a,~b,~l,~p]	2	.6%
Residues in disallowed regions	0	.0%

Number of non-glycine and non-proline residues	332	100.0%
Number of end-residues (excl. Gly and Pro)	2	
Number of glycine residues (shown as triangles)	24	
Number of proline residues	26	

Total number of residues	384	

Based on an analysis of 118 structures of resolution of at least 2.0 Angstroms and R-factor no greater than 20%, a good quality model would be expected to have over 90% in the most favoured regions.