# *In Silico* Homology Modeling and Epitope Prediction of Nucleocapsid Protein region from Japanese Encephalitis Virus

#### A.G. Ingale\*

Department of Biotechnology, School of life sciences, North Maharashtra University, Jalgaon (Maharashtra) 425001, India

#### Abstract

The 3D structure of a protein is a prerequisite for structure based drug design as well as for identifying the conformational epitopes that are essential for the designing vaccines. A 3-dimensional model (3D) was developed for the nucleocapsid protein of Japanese encephalitis virus. A homology modeling method was used for the prediction of the structure. For the modeling, one template proteins were obtained by mGenTHERADER, namely the high-resolution X-ray crystallography structure of NS3 protease helicase of murry vally encephalitis virus(2WV9). By comparing the template protein a rough model was constructed for the target protein using SWISSMODEL, a program for comparative modelling. The model was validated using protein structure checking tools such as Verify3D for reliability. The total of 138 such epitope regions/sites have been identified by kolaskar and tongaokar method. Conformational epitopes are mapped from the 3D structure of nucleocapsid protein of Japanese encephalitis virus modeled using the concept of an antigenic domain. The information thus discussed provides insight to the molecular understanding of nucleocapsid protein of JE virus. The predicted 3-D model may be further used in characterizing the protein in wet laboratory.

**Keywords:** Homology modelling; JE virus; Swiss-PdbViewer; Verify3D

## Introduction

Japanese encephalitis (JE) is an acute viral infection of the central nervous system. Disease normally occur in large numbers in many Asian countries, including Cambodia, China, India, Korea, Myanmar, Philippines, Nepal, Sri Lanka, Thailand and Vietnam. The annual incidence of clinical infection in endemic areas ranges from 10-100 per 100,000 populations. Approximately 3 billion people and 60% of the world's population live in endemic regions and about 50,000 JE cases are notified annually. Mortality happen in JE infection is 10,000. The vast majority of JE infections are inapparent and only one in 25-1000 infections results in symptomatic illness (Umenai et al., 1985). The principal clinical manifestation of illness is encephalitis. JEV infection of host cells produces three structural and seven nonstructural proteins. One of these, the nucleocapsid protein, is the major protein of the virion. This protein is believed to play an important role in a number of processes, including viral attachment, membrane fusion, and entry into the host cell. In response to JEV infection, the host produces virus neutralizing antibodies and cytotoxic T cells (CTLs). The principal target for the neutralizing antibodies is the nucleocapsid protein. It has been shown that protection against JEV infection is mainly antibody dependent, and virusneutralizing antibodies alone are sufficient to impart protection (Konishi et al., 1999; Pan et al., 2001). Peptide(s) from JEV E protein that forms the virus-neutralizing epitope(s) could, therefore, be used for inducing JEV-neutralizing antibodies (Saini and Vrati, 2003)4. The E protein of Japanese encephalitis virus (JEV) is the major antigen used to elicit neutralizing antibody response and protective immunity in hosts (Wu et al., 2003). In the treatment of Japanese encephalitis disease state patients are suffering from various adverse reaction and contraindications also limitation of therapy for the same. One of the major treatments is gene therapy or recombinant DNA vaccines involve targeting multiple antigenic component of virus to direct and empower the immune system to protect the host from viral infection. B-cell epitopes on JEV nucleocapsid

protein are important determinants of protection against virus infection. Peptide based vaccines in which small peptides derived from target proteins epitopes are used to provoke an immune reaction. That has attracted considerable attention recently as a potential means both of treating infectious diseases and promoting the destruction of cancerous cells by a patient's own immune system (Florea et al., 2003). As our knowledge of the immune responses to a protein antigen progressed, it became clear that the whole protein was not necessary for raising the immune response, but small segments of protein called the antigenic determinants or the epitopes were sufficient for eliciting the desired immune response. Its immune protecting capacity is also checked in laboratory to confirm that peptide can be used as potential vaccine. Thus it has been showed that, the use of Bioinformatics tools and techniques not only reduce the time required to identify the candidate peptide as vaccine but also provides an insight in structure function relationship of virus protein.

## Materials and Methods

### Retrieval of target sequence

The amino acid sequence of the nucleocapsid protein of Japanese encephalitis virus was obtained from the sequence database of NCBI http://www.ncbi.nlm.nih.gov/entrez ID: AAA46248. It was ascertained that the three-dimensional structure of the protein was not available in Protein Data

\*Corresponding author: A.G. Ingale, Department of Biotechnology, School of life sciences, North Maharashtra University, Umavi Nagar, PB-80, Jalgaon (Maharashtra) India, Tel: +91-0257-2257423; Fax-+91-0257-2257406; E-mail: agibiotech@gmail.com

Received April 05, 2010; Accepted April 30, 2010; Published April 30, 2010

Citation: A.G. Ingale (2010) *In Silico* Homology Modeling and Epitope Prediction of Nucleocapsid Protein region from Japanese Encephalitis Virus. J Comput Sci Syst Biol 3: 053-058. doi:10.4172/jcsb.1000056

**Copyright:** © 2010 A.G. Ingale. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Bank, hence the present exercise of developing the 3D model of the nucleocapsid protein of Japanese encephalitis virus was undertaken. The protein is 3432 amino acids in length.

# Template searching

An attempt was made to find a suitable template protein for the modeling of the target protein. The template protein was searched through mGeneThrader (Jones, 1999), which is an online tool for searching similar sequences, based on sequence and structure-wise similarity. From the homology searching, one template was selected. Crystallography structure of the NS3 protease helicase of Murry Vally encephalitis virus (2WV9) (Assenberg et al., 2009) was selected as template proteins.

# Sequence alignment

Amino acid sequence alignment of target and template proteins was derived using the Swiss-PdbViewer package (Guex and Peitsch, 1997). Default parameters were applied and the aligned sequences were inspected and adjusted manually to minimize the number of gaps and insertions.

# Homology modeling and structure refinement

A rough 3-D model was constructed from the sequence alignment between nucleocapsid protein of JE virus and the template proteins using SWISSMODEL (Arnold et al., 2006; Schwede et al., 2003) with parameters of energy minimization value. The model was further checked with verify3D (Luthy et al., 1992) and Ramachandran plot at PROCHECK (Laskowski et al., 1993). Accessible surface area prediction using VADAR was performed (Willard et al., 2003). The rough model constructed was solvated and subjected to constraint energy minimization with a harmonic constraint of 100 kJ/mol/Å<sup>2</sup>, applied for all protein atoms, using the steepest descent and conjugate gradient technique to eliminate bad contacts between protein atoms and structural water molecules. Computations were carried out *in vacuo* with the GROMOS96 43B1 parameters set, implementation of SWISS-pdbVIEWER.

# Model validation

The constructed model of nucleocapsid protein of JE virus was examined for validation using different criteria. In the last step of homology modeling the refined structure of the model was subjected to a series of tests for testing its internal consistency and reliability. Backbone conformation was evaluated by the inspection of the Psi/Phi Ramachandran plot obtained from PROCHECK analysis. The Swiss-PdbViewer energy minimization test was applied to check for energy criteria in comparison with the potential of mean force derived from a large set of known protein structures. Packing quality of the refined structure was investigated by the calculation of PROCHECK Quality Control value. The Ramachandran plot of phi/psi distribution in the model is developed using PROCHECK for checking non-GLY residues at the disallowed regions. Standard bond lengths and bond angles of the model were determined using WHAT IF (Vriend, 1990).

## Prediction of antigenic peptides

This program predicts segments from a nucleocapsid protein sequence that are likely to be antigenic by eliciting an antibody response. Antigenic peptides are determined using the method of Kolaskar and Tongaonkar (Kolaskar and Prasad, 1990) (www. mifoundation.org). Predictions are based on a table that reflects the occurrence of amino acid residues in experimentally known segmental epitopes. Segments are only reported if the have a minimum size of 8 residues. The reported accuracy of method is about 75%.

## Solvent accessible regions

Solvent accessible prediction of nucleocapsid protein sequence was done using Vadar. Solvent accessible scales for delineating hydrophobic and hydrophilic characteristics of amino acids and scales are developed for predicting potential antigenic sites of globular proteins, which are likely to be rich in charged and polar residues. It was shown that a nucleocapsid protein JE virus is hydrophobic in nature and contains segments of low complexity and high-predicted flexibility.

# MHC binding peptide prediction

These MHC binding peptides are sufficient for eliciting the desired immune response. The prediction is based on cascade support vector machine, using sequence and properties of the amino acids. The correlation coefficient of 0.88 was obtained by using jack-knife validation test.

# **Results and Discussion**

# Structural description of the model

The three-dimensional (3D) structure details of proteins are of major importance in providing insights into their molecular functions. Further analysis of 3D structures will help in the identification of binding sites and may lead to the designing of new drugs. The protein sequence of the of nucleocapsid protein of JE virus was obtained from the NCBI sequence database. Multiple alignment of the primary structure of the target protein highlights the degree of sequence conservation and high sequence similarity. Homology modeling is only a viable technique because it produces models that can be used for further research. Homology modeling helps in predicting the 3-D structure of a macromolecule with unknown structure (target) by comparing it with a known template from another, structurally highly similar, macromolecule. The structure of the target protein is structurally similar with the template if both the target and template sequences are similar. In general, 30% sequence homology is required for generating useful models. Here, the sequence alignment score was 44 as calculated by ClustalW. In our study, based on the results obtained from mGenTHREADER program, the X-ray structure of the NS3 protease helicase of murry vally encephalitis virus (2WV9) were selected as templates. The three-dimensional structure of the Egp of JEV was predicted using the knowledge-based homology modeling approach and X-ray structure data of the Egp of tickborne encephalitis virus as a template (Kolaskar and Kulkarni-Kale, 1999). SWISSMODEL was used for building the model and global energy minimization. The sequence was obtained from sequence database and was submitted to blastp search. After the BLAST analysis, PROCHECK was used to validate the model.

The predicted structure was refined by solvating the protein in a 10-Å layer of water by explicitly considering 4867 water molecules. Four independent structure evaluation methods report this structure to be acceptable stereochemically and geometrically. The nucleocapsid protein of JEV has an extended



**Figure 1:** (A) Predicted 3-D structure of Nucleocapsid protein of JE virus :Ramachandran plot analysis. The Plot statistics are: residues in most favoured regions [A,B,L] - 104 (85.0%); residues in additional allowed regions [a,b,l.p] - 72 (14.1%); residues in generously allowed regions [-a,-b,-l,-p] - 0 (0.0%); residues in disallowed regions – 3(0.6%); number of non-glycine and non-proline residues- 511 (100.0%); number of end residues (excl. Gly and Pro) - 2; number of glycine residues (shown as triangles) - 55; number of proline residues - 33; total number of residues-601. (B) Predicted 3-D structure of Nucleocapsid protein of JE virus.

structure with forty two b-sheets, twenty two a-helices, and four domains. The water-solvated structure was used to delineate conformational and sequential epitopes. These results document the importance of tertiary structure in understanding the antigenic properties of flaviviruses in general and JEV in particular. The conformational epitope prediction method could be used to identify conformational epitopes on any protein

antigen with known three-dimensional structure. This is one of the largest proteins whose three-dimensional structure has been predicted using a homology modeling approach and water as a solvent.

The refined model was analyzed by different protein analysis programs including PROCHECK for the evaluation of the

Citation: A.G. Ingale (2010) In Silico Homology Modeling and Epitope Prediction of Nucleocapsid Protein region from Japanese Encephalitis Virus. J Comput Sci Syst Biol 3: 053-058. doi:10.4172/jcsb.1000056



Ramachandran plot quality, and WHATIF for the calculation of packing quality. The verify 3D plot show the 81.89 % of residue had an average 3D-1D score 70.2 (Figure 2). This structure for the corresponding coordinates in pdb format) was found to be satisfactory based on the above results. The predicted 3-D model of the nucleocapsid protein of JE virus will be very useful in wet laboratory while studying the real structure of the protein. The total energy values of the predicted 3-D model were calculated as 86.0% of Ramachandran plot value in 30 and 40 steepest descent and conjugate gradient, respectively. The protein model has more b-sheets with only three short a-helices. On comparison with the model of the Egp of Japanese Encephalitis virus (Kolaskar and Kulkarni-Kale, 1999), it is found that there is an additional a - helix that may play an important role and may elicit certain function. The same observation is made with the two template structures. The model and templates are found to have similar folds.

#### B cell epitope

JEV infection in a host produces both the virus-neutralizing antibodies. It has, however, been shown that protection against JEV is mainly antibody dependent and that virus-neutralizing anti-bodies alone are sufficient to impart protection. This is also evident from the fact that the formalin-inactivated JEV vaccine, which does not induce CTLs (Halstead, 1998), provides protection to vaccines against JEV. It therefore appears that the B-cell epitopes on JEV nucleocapsid protein are important

Position	Epitope	Score
1	MTKKPGGPGKNRAINMLKRG	1
1817	IFMTATPPGTTDPFPDSNAP	1
1667	SAIVQGDRQEEPVPEAYTPN	1
2525	LKRGRPGGRTLGEQWKEKLN	1
2882	RVFKEKVDTKAPEPPAGAKE	1
878	VDLSVVVNKPVGRYRSAPKR	0.998
2669	TLFCDIGESSPSPEVEEQRT	0.998
438	HGTTTSENHGNYSAQVGASQ	0.998
1936	PTILEEGEGRVILGNPSPIT	0.998
3053	GKQGGKMYADDTAGWDTRIT	0.998
2985	GKREKKPGEFGKAKGSRAIW	0.998
1510	DTPSPKPCSKGDTTTGVYRI	0.997
1633	DYPRGTSGSPILDSNGDIIG	0.997
2480	TLWDNGASAVWNSTTATGLC	0.996
1961	QRRGRVGRNPNQVGDEYHYG	0.995
186	GNDPEDVDCWCDNQEVYVQY	0.992
2797	GEVHSNQEKIKKRIQKLKEE	0.988
1323	IGICSLLHERKKTMAKKKGA	0.985

 Table 1: BCPred Predictions of Nucleocapsid protein of JE virus.

determinants of protection against virus infection. Predicted B cell epitopes are presented in (Table 1). Following are the B cell epitopes having high score position(1) 'MTKKPGGPGKNR AINMLKRG'(1817)'IFMTATPPGTTDPFPDSNAP'(1667)' IFMTA TPPGTTDPFPDSNAP'(2525)'LKRGRPG GRTLGEQWKEKLN'. As our knowledge of the immune responses to a protein antigen progressed, it became clear that the whole protein was not necessary for raising the immune response, but small segments of protein called the antigenic determinants or the epitopes were sufficient for eliciting the desired immune response.

#### Prediction of antigenic peptides

Nucleocapsid protein (Japanese encephalitis) Sequence is 3432 residues long. There are 138 antigenic determinants in sequence (Figure 3a). The highest pick at start position - 2254 to 2272. The sequence is- NQLAVFLICVLTVVGVVAA. The average for the whole protein is above 1.0 then all residues having above 1.0 are potentially antigenic. The highest pick sequence of antigenic determinant site is used for insertion at the multiple cloning vectors. Highest pick in antigenic determinants plot indicate antigenic site for the host cell attachments. If this sequence is inserted into the multiple cloning site of cloning vectors, gives new synthetic clone vector, which is used as synthetic peptide vaccines. Average antigenic propensity for this protein is 1.0249. The ability of an individual antibodycombining site to react with only one antigenic determinant. The ability of a population of antibody molecules to react with only one antigen.

To improve immunogenicity, the peptides could be chemically conjugated to a large carrier protein. However, the process of chemical conjugation is not very reproducible, and uniformity of the peptide density on the carrier protein cannot be ensured. Recent findings show that peptides presented in a particulate form result in enhanced immune responses (Lomonossoff and Johnson, 1996). The conformation of the peptide is an important determinant of its immunogenicity, and it may determine whether the anti-peptide antibodies would also recognize the native protein from which the peptide was derived. Thus, to improve chances of producing anti-peptide antibodies capable of recognizing JEV nucleocapsid protein. Reverse Translation of Antigenic site sequence. It is method for design a DNA sequence that code for protein. It is useful site for the Peptide antigen design principles, can be used to help ensure production of successful antibodies.



# The promiscuous MHC binding peptide prediction

In this test, we found the MHC 1 and MHC 2 binding regions. MHC molecules are cell surface glycoproteins, which take active part in host immune reactions and involvement of MHC class 1 and MHC 2 in response to almost all antigens (Gibbons et al., 2004). In this assay we predicted the binding affinity of nucleocapsid protein of JE virus having 68 amino acids, which shows 60 different nonamers.For development of MHC binder prediction method, an elegant machine learning technique supports vector machine (SVM) has been used. SVM has been

trained on the binary input of single amino acid sequence. MHC2Pred predicts peptide binders to MHCI and MHCII molecules from protein sequences or sequence alignments using Position Specific Scoring Matrices (PSSMs). In addition, we predict those MHCI ligands which C-terminal end is likely to be the result of proteosomal cleavage (Nielsen et al., 2005; Kesmir et al., 2002). The MHC peptide binding is predicted using neural networks trained on C terminals of known epitopes. In analysis predicted MHC/peptide binding is a log-transformed value related to the IC50 values in nM units. These MHC binding peptides are sufficient for eliciting the desired immune

response. The predicted binding affinity is normalized by the 1% fractil. The MHC peptide binding is predicted using neural networks trained on C terminals of known epitopes (Bhasin and Raghava, 2005).

## Conclusion

Developing successful vaccine for Japanese encephalitis will likely involved targeting multiple antigenic component of virus to direct and empower the immune system to protect the host from viral infection. DNA immunization hold grate promise for providing safe and inexpensive vaccine for many infectious pathogens, including Japanese encephalitis. The direct injection of foreign genes by genetic immunization has resulted in specific immune responses that exhibit characteristic of protective immunity against the number of infectious agent in small animal model as well as primate studies. Genetic vaccination cassettes targeting each of Japanese encephalitis genes. Such a collection of immunization cassettes should be capable of immunizing of broad immunity against both humoral and cellular epitope thus giving vaccines the maximum ability to deal with viral immune escape. Further more, the power of the genetic adjuvant coadminatration (consisting co-stimulatory molecule or cytokine genes) to dramatically direct and enhance the antigen - specific humoral and cellular immune responses induce by DNA immunogens to analogous to the potential demonstrated by the combination therapy consisting of three different antiflaviviral agent for the control of Japanese Encephalitis pathogenesis. Therefore, it is important to establish the structure-function relation of nucleocapsid protein of JE virus. However, a solved structure for nucleocapsid protein of JE virus is not available at the protein data bank (PDB). Therefore, we created a model of envelope glycoprotein of nucleocapsid protein of JE virus using the X-ray structure of the NS3 protease helicase of murry vally encephalitis virus (2WV9) as template with swissmodel (a comparative modeling program). The model was validated using protein structure checking tools such as PROCHECK, WHAT IF and Verify 3D for reliability. Thus, to improve chances of producing anti-peptide antibodies capable of recognizing JEV nucleocapsid protein. We theorized that a successful immunization strategy against Japanese Encephalitis infection could also involved a peptide vaccine.

#### References

- Arnold K, Bordoli L, Kopp J, Schwede T (2006) The SWISS-MODEL Workspace: A web-based environment for protein structure homology modelling. Bioinformatics 22: 195-201. » CrossRef » PubMed » Google Scholar
- Assenberg R, Mastrangelo E, Walter TS, Verma A, Milani M, et al. (2009) Crystal Structure of a Novel Conformational State of the Flavivirus Ns3 Protein: Implications for Polyprotein Processing and Viral Replication. J Virol 83: 12895-12906. » CrossRef » PubMed » Google Scholar
- Bhasin M, Raghava GPS (2005) Pcleavage: an SVM based method for prediction of constitutive proteasome and immunoproteasome cleavage sites in antigenic sequences. Nucleic Acids Res 33: W202-207. » CrossRef » PubMed » Google Scholar
- Florea L, Halldórsson B, Kohlbacher O, Schwartz R, Hoffman S, et al. (2003) Epitope Prediction Algorithms for Peptide-based Vaccine Design. Proc IEEE

Comput Soc Bioinform Conf 2: 17-26. » CrossRef » PubMed » Google Scholar

- Gibbons DL, Vaney MC, Roussel A, Vigouroux A, Reilly B, et al. (2004) Conformational change and protein-protein interactions of the fusion protein of Semliki Forest virus. Nature 427: 320-325. » CrossRef » PubMed » Google Scholar
- Guex N, Peitsch MC (1997) SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modelling. Electrophoresis 18: 2714-2723. » CrossRef » PubMed » Google Scholar
- Halstead SB (1998) Pathogenesis of dengue: challenges to molecular biology. Science 239: 476-481. » CrossRef » PubMed » Google Scholar
- Jones DT (1999) mGenTHREADER: an efficient and reliable protein fold recognition method for genomic sequences. J Mol Biol 287: 797-815. » CrossRef » PubMed » Google Scholar
- Kesmir C, Nussbaum A, Hansjorg S, Vincent D, Brunak S (2002) Prediction of proteasome cleavage motifs by neural networks. Prot Eng 15: 287-296.
   » CrossRef » PubMed » Google Scholar
- Kolaskar AS, Kulkarni-Kale U (1999) Prediction of Three-Dimensional Structure and Mapping of Conformational Epitopes Of Envelope Glycoprotein of Japanese Encephalitis Virus. Virology 261: 31-42. » CrossRef » PubMed » Google Scholar
- Kolaskar AS, Prasad TA (1990) Semi-empirical method for prediction of antigenic determinants on protein antigens. FEBS Letters 276: 172-174.
   » CrossRef » PubMed » Google Scholar
- 12. Konishi E, Yamaoka M, Khin-Sane-Win, Kurane I, Takada K, et al. (1999) The Anamnestic Neutralizing Antibody Response Is Critical for Protection of Mice from Challenge following Vaccination with a Plasmid Encoding the Japanese Encephalitis Virus Premembrane and Envelope Genes. J Virol 73: 5527-5534. » CrossRef » PubMed » Google Scholar
- Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK a program to check the stereochemical quality of protein structures. J App Cryst 26: 283-291. » CrossRef » PubMed » Google Scholar
- 14. Lomonossoff GP, Johnson JE (1996) Development of cowpea mosaic virus as Vaccine. Intervirology 14: 799-810. » CrossRef » PubMed » Google Scholar
- Luthy R, James UB, Eisenberg D (1992) Assessment of protein models with three-dimensional profiles. Nature 356: 83-85. » CrossRef » PubMed » Google Scholar
- 16. Nielsen M, Lundegaard C, Lund O, Kesmir C (2005) The role of the proteasome in generating cytotoxic T cell epitopes: Insights obtained from improved predictions of proteasomal cleavage. Immunogenetics 57: 33-41. » CrossRef » PubMed » Google Scholar
- 17. Pan CH, Chen HW, Huang HW, Tao MH (2001) Protective Mechanisms Induced by a Japanese Encephalitis. J Immunol 166: 7419 - 7426. » CrossRef » PubMed » Google Scholar
- 18. Saini M, Vrati S (2003) A Japanese encephalitis virus peptide present on Johnson grass mosaic virus-like particles induces virus-neutralizing antibodies and protects mice against lethal challenge. J Virol 77: 3487-94. » CrossRef » PubMed » Google Scholar
- Schwede T, Kopp J, Guex N, Peitsch MC (2003) SWISS-MODEL: an automated protein homology-modeling server. Nucliec Acid Research 31: 3381-3385.
   » CrossRef » PubMed » Google Scholar
- Umenai T, Krzysko R, Bektimirov TA, Assaad FA (1985) Japanese encephalitis: current worldwide status. Bull WHO 63: 625-631. » CrossRef » PubMed » Google Scholar
- 21. Vriend G (1990) WHAT IF: A molecular modeling and drug design program. J Mol Graph 8: 52–55. » CrossRef » PubMed » Google Scholar
- 22. Willard L, Ranjan A, Zhang H, Monzavi H, Robert FB, et al. (2003) Wishart. VADAR: a web server for quantitative evaluation of protein structure quality. Nucleic Acids Res 31: 3316-3319. » CrossRef » PubMed » Google Scholar
- 23. Wu SC, Yu CH, Lin CW, Chu IM (2003) The domain III fragment of Japanese encephalitis virus envelope protein: mouse immunogenicity and liposome adjuvanticity. Vaccine 21: 2516-22. » CrossRef » PubMed » Google Scholar