Prediction of 3D Structure of Envelope Glycoprotein of Sri Lanka Strain of Japanese Encephalitis Virus

Urmila Kulkarni-Kale & A. S. Kolaskar

Bioinformatics Centre, University of Pune, Pune 411 007. India.

urmila@bioinfo.ernet.in; kolaskar@bioinfo.ernet.in

Abstract

This paper describes knowledge-based homology modeling studies of envelope glycoprotein (Egp) of Sri Lanka strain of Japanese encephalitis virus (JEVS). JEVS is a mosquito-borne *Flavivirus*, which is an important human pathogen. The Egp is a major structural antigen and is responsible for viral hemagglutination and neutralisation. The 3D structure of 399 amino acids from the extra cellular domain of Egp of JEVS has been predicted using the x-ray crystal structure of Egp of Tick-borne encephalitis virus as a template and the knowledge-based homology modeling approach. Even though the homology modeling is the best method for prediction of 3D structure, prediction of structures of loop regions is still a challenge. A novel approach of molecular dynamics simulations and geometry optimisation has been used to sample the conformations of loop regions. The Egp of JEVS has an extended structure with nine β-sheets, two α-helices and three domains. The predicted structure was compared with the model of Egp of Nakayama strain of Japanese encephalitis virus (JEVN), which was developed earlier (Kolaskar & Kulkarni-Kale, 1999). Similarities and differences between the structures of Egps of two strains of JEV are discussed. These models illustrate effect of mutations on the local and global conformation of Egp and help to explain strain specific properties. The sequential and conformational epitopes of Egp of JEV were predicted using an algorithm developed in house (Kulkarni-Kale, 2002). The predicted B cell epitopes could be used to design synthetic peptide vaccine against JEV.

Keywords: Japanese encephalitis virus, envelope glycoprotein, three-dimensional structure, homology modeling, molecular dynamics simulations, loop modeling, mutations, strain specific property, conformational epitope, sequential epitope, antigenic determinant, peptide vaccine.

1 Introduction

The field of Bioinformatics is growing at a phenomenal rate and the task in the post-genomic era is to prove how *in-silico* simulations facilitate wet lab experiments. It is known that the protein structure is responsible for the function and is conserved than the sequence. While the structural genomics speeds up the rate of structure determination (Burley & Bonanno, 2002; Liu & Rost, 2002), the computational approaches like homology modeling can expand that knowledge to large number of protein families to fill the gap between number of known sequences and structures (Baker & Sali, 2001; Peitsch, 2002). The reliable models obtained using homology-modeling method could be used to understand protein function and to design strategies for development of drugs and vaccines.

Japanese encephalitis (JE) is a mosquito-borne arboviral disease of major public health importance in Southeast Asia including India (Banerjee, 1996**;** Monath & Heinz, 1996; Chambers, et. al., 1997). It is the principal cause of viral encephalitis worldwide (Tsai, 2000). Recently, JE outbreaks have also been reported in Australia (Van Den Hurk, et al., 2001; McCormack & Allworth, 2002), Malaysia (Easton, 1999), Nepal and Taiwan (Monath, 2002). In Southeast Asia, JE was first reported in Sri Lanka (Monath & Heinz, 1996; Tsai, 2000).

The causative agent of JE is Japanese encephalitis virus (JEV), which belongs to family *Flaviviridae*. Flaviviruses are enveloped, positive sense, single stranded RNA viruses and are spherical in shape with a diameter of ~40 nm. The genome is ~11kb, which synthesizes a polyprotein containing capsid (C), membrane (M), envelope glycoprotein (Egp) and non-structural proteins such as NS1, NS2A, NS2B, NS3, NS4A, NS4B and NS5. The Dengue type 1, Dengue type 2, Dengue type 3, Looping ill virus, Murray Valley encephalitis virus, West Nile virus, Yellow fever virus and Tick-borne encephalitis virus are some of the important pathogenic viruses that belong to the Flavivirus group (Monath & Heinz, 1996; Van Regenmortel, et al., 2000).

JEV maximally affects children in the age group of 5-15 years and adults in the age group of 40-45 years. More than 35,000 cases and 10,000 deaths have been reported annually due to JE (Ellis, et al., 2000, Shlim & Solomon, 2002; Daley & Dwyer, 2002). The incubation period of JEV is 5-15 days. Illness usually begins with an abrupt onset of high fever associated with change in mental status followed by motor dysfunctions. The encephalitis in children usually begins with gastrointestinal symptoms associated with progressive decline in alertness eventually leading to coma of varying degree. Approximately 25% of the cases are fatal, children (under 10 years) are more likely to die, and if they survive, they are likely to have residual neurological deficits Overall, approximately one third of surviving patients exhibit serious residual neurological disability (Tsai, 2000).

Currently, mouse brain-derived, inactivated JE vaccine is available internationally (Monath, 2002), which is stable at 4ºC. After reconstitution, vaccine is stable at 22ºC for at least two weeks, but at 37ºC, potency declines to 85% (Gowal, et al., 1991). Furthermore, the protective immunity is introduced only after administration of

multiple doses (Hoke, et al., 1988; Mohan Rao, et. al., 1993; Gambel, et al., 1995). Thus, the vaccine requires the cold chain for storage and transport and the cost of immunization becomes high in tropical countries, where JEV is endemic (Chambers, et al., 1997). In view of the occurrences of JE epidemics worldwide, the WHO has placed a high priority on the development of a new vaccine for prevention of JE (Monath, 2002). Therefore, it becomes essential to explore new approaches such as DNA or peptide vaccines for development of vaccines against JEV.

In case of JE virus and closely related flaviviruses, it has been demonstrated that the type specific, neutralizing and protective antibody response is observed against Egp (Venugopal & Gould, 1994; Chambers, et. al., 1997). These observations suggest that synthetic peptides derived from Egp, could be used to induce protective immune response against JE virus infection. The idea of using peptides as vaccine is based on the fundamental principle of immunology that the immune system recognizes only specific regions on the antigenic proteins called antigenic determinants or epitopes. Thus, logically, introducing a peptide representing B-cell and or T-cell epitope should induce a similar response as an antigen having these epitope (Wisdom, 1992; Dyson & Wright, 1995). This simple and seemingly convincing concept has generated a lot of interest and the efforts are on to develop peptide vaccines against bacteria, viruses, parasites and also for diseases like cancer (Weber, 2002; Offringa et al., 2000). One of the lacunae in designing the peptide vaccines has been lack of three-dimensional (3D) structure information of the antigenic proteins and an algorithm to predict B-cell epitopes. We have earlier predicted the 3D structure of JEVN (Nakayama strain of JEV) and developed a method to predict the sequential and conformational epitopes using accessibility data in an explicit fashion (Kulkarni-Kale, 2002, Kolaskar & Kulkarni-Kale, 1999). In this paper, we report the prediction of 3D structure of Sri Lanka strain of JEV using homology modeling approach. There are 8 mutations in the Egp of JEVS (EMBL: Z34097) as compared to JEVN (PIR-PSD: GNWVJE). The predicted structures of Egps of JEVS and JEVN not only help to understand the effect of mutations on the local and global conformation of Egp but also explain the strain specific biological properties. This knowledge, if employed in the selection of candidate peptide would help in development of synthetic peptide vaccine against JEV.

2 Materials & Methods

2.1 Materials

The sequence of envelope glycoprotein (Egp) of Sri Lanka strain (691004) of Japanese encephalitis virus (Paranjpe & Banerjee, 1996) was obtained from the EMBL database (Acc. No: Z34097).

The X-ray structure data of the Egp of Tick-borne encephalitis virus (TBEV) solved at 1.9Å resolution (Rey et al., 1995) was used as the template structure (PDB entry: 1SVB).

The InsightII® molecular modeling package (version: 95, MSI and version 2000, Accelrys) and the Homology® module was used to build and optimise the model.

The molecular dynamics simulations of the loop regions were carried out using Discover® module and the Analysis® module was used to analyse the trajectory data.

The predicted structure was critically evaluated using five independent methods such as PROSTAT (module in Homology), MM and MX (programs developed in-house, Kolaskar & Choudhary, unpublished), ProsaII (Sippl, 1993), PROCHECK (Laskowski, et. al., 1993) and WHAT_CHECK program from WHAT_IF suite (Vriend, 1990).

Interactomic contacts were analysed using CSU software (Sobolev et al., 1999).

The antigenic determinants and conformational epitopes were predicted using the algorithm developed in-house (Kulkarni-Kale, 2002, Kolaskar & Kulkarni-Kale, 1999).

2.2 Methods

2.2.1 Building Homology Model

The model of Egp of JEVS was built independently by using the structure of Egp of TBEV as a template rather than using the model of JEVN, which was developed earlier (Kolaskar & Kulkarni-Kale, 1999). However, the model of Egp of JEVS was built and optimised using the same parameters and methodology that was used to predict 3D structure of Egp of JEVN, which has been reported earlier.

2.2.2 Prediction of the structures of loops using molecular dynamics (MD) approach

The structures of the loop regions were predicted using a novel approach of the molecular dynamics simulations. The initial conformations of the residues in the loop region were perturbed to reduce the conformational bias and to sample the optimum conformation for every loop. The MD was carried out at 300K for 500ps. The equilibration was carried out for 100ps. This was followed by the conformational sampling for 500ps. The conformers were captured at interval of 10ps and optimised using steepest descents and conjugate gradients methods till the average rms derivative reached 0.01 and 0.001 Kcal/mol/Å respectively.

The trajectory data thus generated for every loop was analysed in detail for both, the initialisation and production phase. The conformer with the minimum most energy and allowed conformation was used to assign the structure to the respective loop region.

3 Results & Discussions

3.1 Model building & refinement

The structure of the Egp of Sri Lanka strain of Japanese encephalitis virus (JEVS) has been modeled independently by using the structure of Egp of Tick-borne encephalitis virus (TBEV) as a template and not by using

the model of Egp of Nakayama strain of Japanese encephalitis virus (JEVN). This was done mainly to avoid any bias that might have been introduced because of the amino acid composition of Egp of JEVN or choices made at various stages of optimisation of the model of Egp of JEVN. The over-all identity in the sequences of Egps of TBEV and JEVS is 40% and as expected all cysteine residues are conserved (Figure 1).

There are eight mutations in the Egp of JEVS when compared to Egp of JEVN. They are A54T, E83K, G153W, I176V, H219R, W225L, K297Q, T321I. Out of these, one mutation (E83K) is in the loop region, the mutation (G153W) is in the hyper variable region of flaviviruses (149-164) and remaining mutations are part of SCRs. These mutations had no effect on the global alignment of Egps of JEVS and TBEV when compared to the alignment of Egps of JEVN and TBEV. This led to similar definitions of SCRs, loop regions (Figure 1) and initial structure for the backbone of the model of Egp of JEVS and JEVN, which was predicted earlier (Kolaskar & Kulkarni-Kale, 1999).

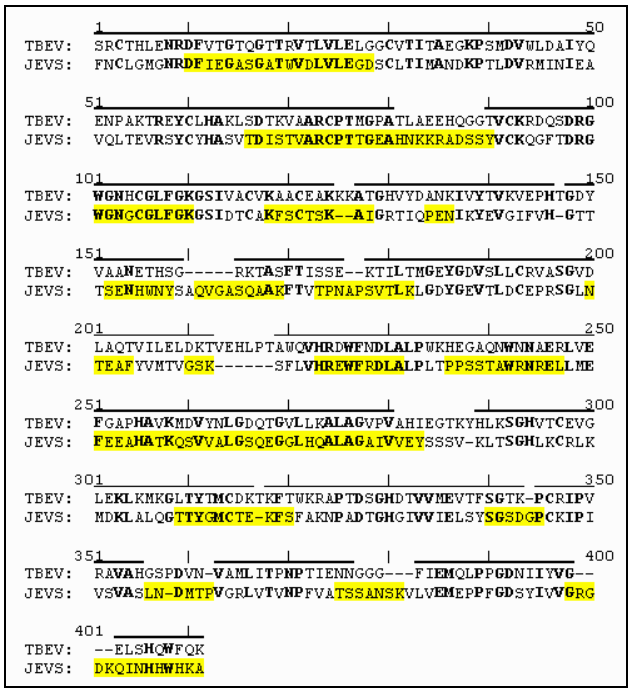

Figure1: Sequence alignment of Egps of TBEV and JEVS obtained using the homology module of InsightII and CSW matrix (Kolaskar & Kulkarni-Kale, 1992). Identical amino acids are shown in bold, the lines on the top of alignment mark SCRs. The numbers on the top shows the alignment positions. Predicted antigenic determinants of Egp of JEVS (Kulkarni-Kale, 2002, Kolaskar & Kulkarni-Kale, 1999) are highlighted.

3.2 Prediction of structures of loops of JEVS using molecular dynamics (MD) approach

Even though the knowledge-based homology modeling is the best method available to predict the 3D structures of proteins, the accuracy and quality of the model depends on two factors, namely the sequence alignment of query and template proteins and the structures of loop regions. Prediction of the structure of the loop regions is the most important and challenging task in building models of proteins using homology-modeling approach as conformations of these regions are sensitive and give rise to inaccuracies in the modeled structures (Johnson, et. al., 1994). These inaccuracies are mainly due to the fact that the loops are often variable in length, sequence and conformation even among the proteins of the same family. Furthermore, being situated at the molecular surface and not in the hydrophobic core region, loops are also found to undergo large movements about the mean positions (Moult, et. al., 1997; Venclovas, et al., 2001; Mart-Renom, et al., 2002). A hybrid approach of molecular dynamics (MD) and energy minimisation was used to perturb the initial loop conformations and to perform conformational search to predict conformations of the loop regions. Molecular dynamics (MD) is a computational method for simulation of motion of particles within a system using Newton's laws of motion. Application of MD techniques to biological macromolecules has yielded information on structural movements (Berendsen & Hayward, 2000). It is also being used in refinement of 3D structures that are determined using the X-ray crystallography and NMR.

The constant-temperature, constant-volume (NVT) ensemble, which is also referred as the canonical ensemble, was used to perform the conformational search for every loop region. The ensemble was obtained by controlling the temperature through direct temperature scaling during the initialisation phase and by temperature-bath coupling during the data collection phase. The volume was kept constant throughout the run. NVT is the most appropriate choice, when the conformational searches of the molecules are performed in vacuum without the periodic boundary conditions. MD simulations were carried out at 300K for 500ps with equilibration for 100ps. The trajectories were sampled at every 10ps and therefore 50 frames have been archived during the MD simulations for 500ps. The MD trajectories of the respective loops were analysed in detail for both, the equilibration phase and the production phase. The temperature and total energy of the conformers was plotted as a function of time to confirm if the system is equilibrated before production phase. One of the methods for assessing the exploration of conformational space in the production phase is to calculate the rmsd of Cα atoms for every structure that has been sampled during MD simulation, from every other structure (Doniach & Eastman, 1999). The plots of Cα distances of conformers sampled in the MD simulation of every loop were drawn to visualise the extent of variation in the sampled conformers. The resulting matrix of deviations was then plotted as a 2D image. The metastable state, in which the system remains for extended period of time show up as squares along the diagonal where as the patches away from diagonal indicate that the system has moved away from a given conformation and returned to it again. The geometrically allowed and minimum energy conformer was thus identified and its conformation was then assigned to respective loop. Thus, structures of 6 loops in Egp of JEVS predicted using MD simulations.

3.3 Analyses of MD trajectories of Loops of JEVS

The Loop1 is 11 amino acid long region between residues 81-91. One out of 8 mutations (E83K) between JEVN and JEVS is located in Loop1. Figure 2 depicts the plots of temperature and total energy against the elapsed time in the equilibration phase. The elapsed time is shown on the X-axis from 100fs to 100ps. This Figure shows that the system has attended the temperature of 300K and was stabilised in the equilibration phase.

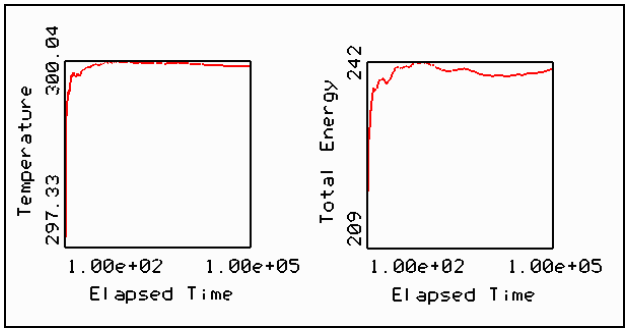

Figure 2: Analysis of equilibration phase of Loop1.

The 50 conformers sampled during the production phase of the MD runs of the Loop1 were analysed to check if the duration of MD simulations was enough to explore the conformational space. The Cα cluster graph for the various conformers that were sampled in the molecular dynamics simulation of Loop1 are shown in Figure 3. The all-against-all rmsd values for Cα atoms of 11 residues from 50 conformers of Loop1 indicate maximum deviation of 3.02Å. The plot of energy of respective conformers is also shown in the same Figure and it depicts various energy basins along the path of MD simulation of Loop1. The sampled conformers are also shown in Figure 3. Two conformers which were captured at 140th ps (frame:14) and 160th ps (frame: 16) are very different as compared to other frames. These frames are shown in yellow and red thin lines, which were found to adopt entirely different conformations, especially around a single residue, which is R85. Figure 4 shows the variation of the dihedral angles (ϕ,ψ) as a function of time. The frames are plotted on the X-axis and both the dihedral angles, ϕ and ψ are plotted on Y-axis. As expected, in the constrained dynamics, the variation in the (ϕ,ψ) angles of the residues at both the ends of the loop is restricted. However, the residues in the middle region of the loop explore the conformational space in an unrestricted fashion, which can be seen from the variation of the dihedral angles, ϕ and ψ of individual residues of Loop1.

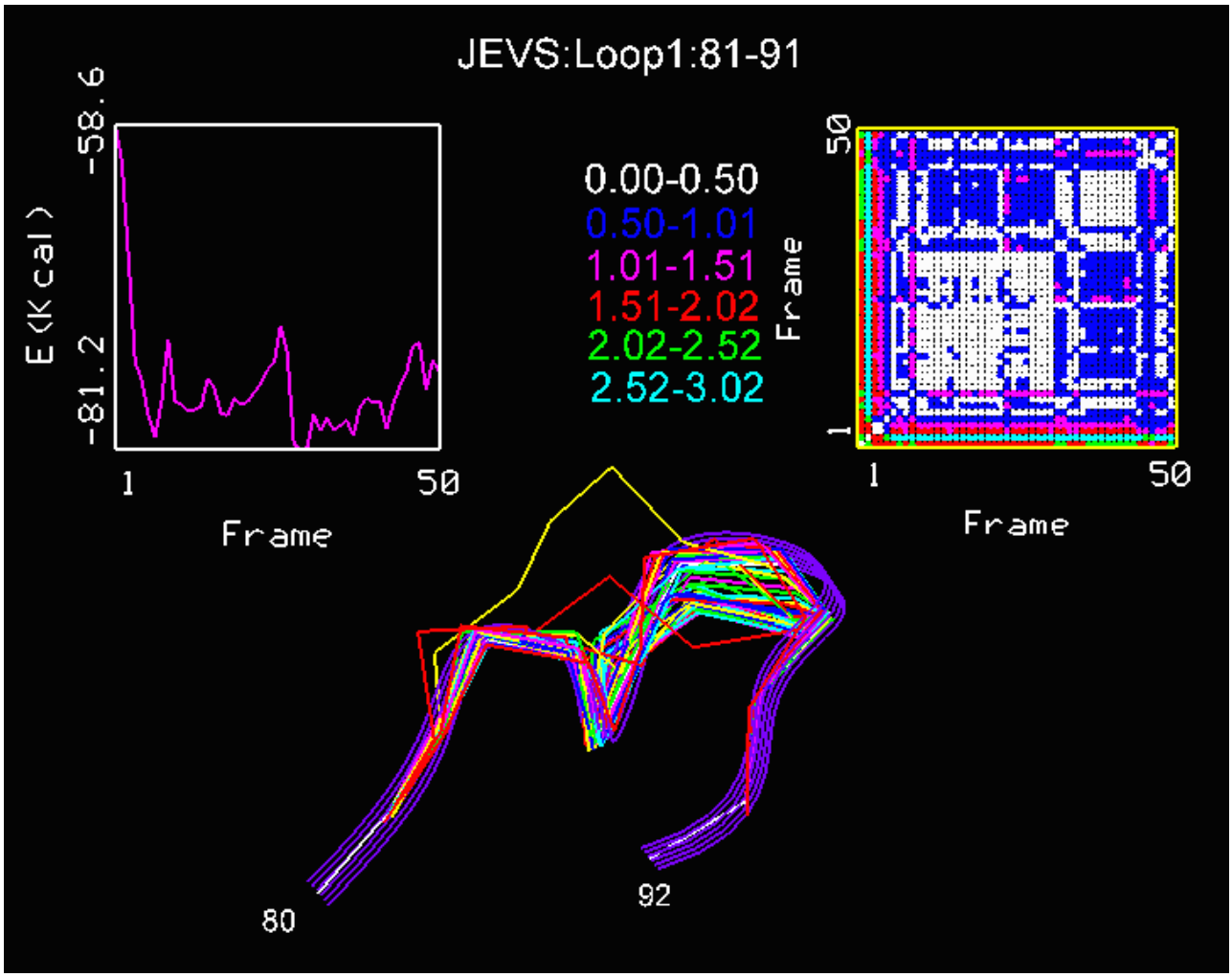

Figure 3: Energy plot, cluster graph and ensemble of conformers of Loop1 of Egp of JEVS. The cluster graph is colour coded as shown in the figure.

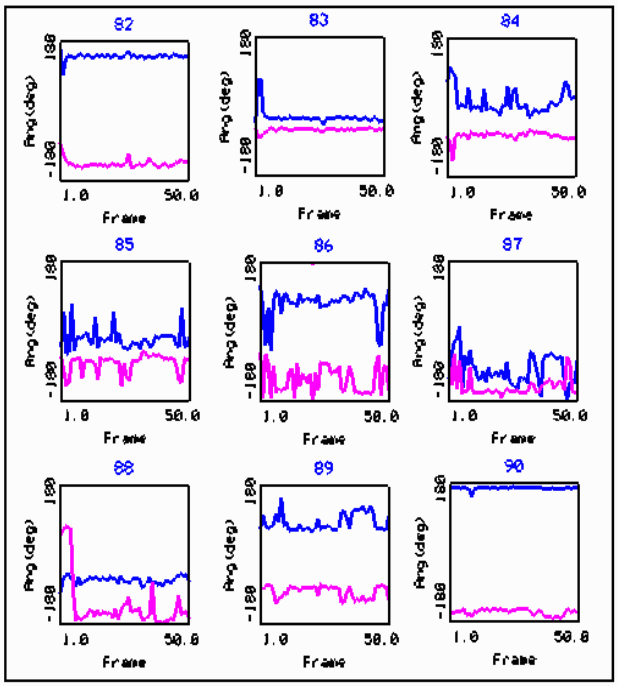

Figure 4: Variation of dihedral angles, ϕ (magenta) and ψ (blue) of individual residues of Loop1 of JEVS. The MD trajectory data after minimisation has been used to plot these graphs.

In all seven conformers, captured as frames 29, 30, 28, 37 and 7 were energetically favourable (-58.8 to –57.6 Kcal) and the 30th conformer was found to be acceptable geometrically. The 29th conformer, even though had minimum most energy had one residue for which the dihedral angles were just outside of the boundaries of the Ramachandran plot (data not shown). Therefore, the conformation of 30th frame having energy of -58.5Kcal was assigned to Loop1 of JEVS.

The analysis of the 500ps dynamics simulation of the Loop2 is shown in the Figure 5. The energy profile of various frames shows that the energy of Loop2 fluctuates between –1.87 to –7.23 Kcal. As can be seen from the ensemble and the rmsd plot, this loop has a highly flexible structure, which essentially deviates around a mean position. The rms deviation of the sampled conformers of Loop2 of JEVS was observed to be 2.74 Å. Loop2 was found to adopt the type I β-turn in JEVS. The flexibility of the Loop2 in case of JEVS could be attributed to the mutation at position 153, where Gly in JEVN is substituted for Trp in JEVS. The conformation of the first frame was assigned to the Loop2 of JEVS as it satisfied all the criteria.

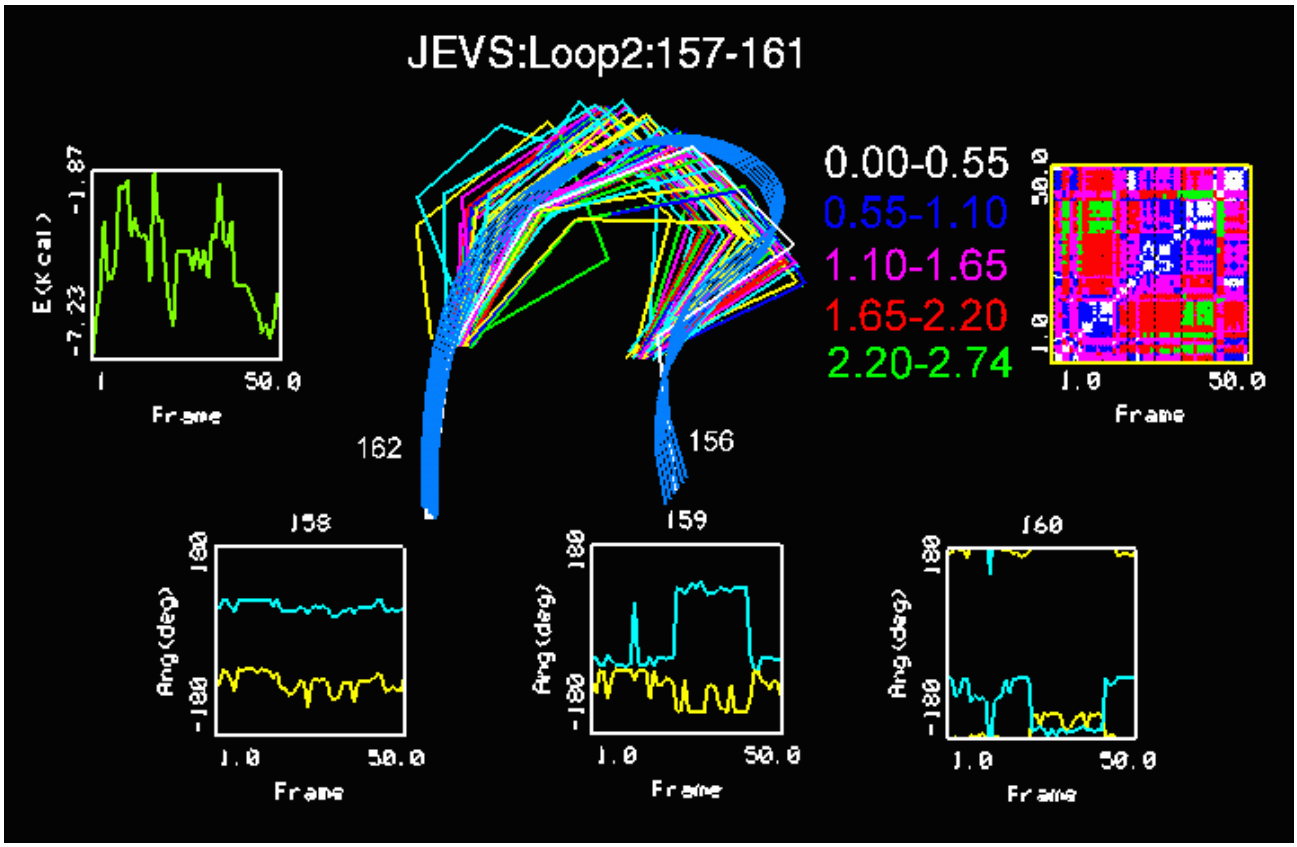

Figure 5: A composite picture of analysis of 500ps MD trajectory of Loop2 of JEVS. The energy plot and the cluster graph are shown along with the (ϕ,ψ) variation plots of individual residues (ϕ : yellow; ψ : cyan) Note that the (ϕ,ψ) plots of the first and the last residue of the Loop2 are not shown. The ensemble of conformers (multi-coloured thin lines) is shown with the ribbon rendering (blue) of the 1st frame (white), the conformation of which was assigned to Loop2.

The Loop3 (272-278) is 7 residues long region. Out of 50 conformers that were sampled, the conformers 17-18, 27-32 and 50 had minimum most energy (-44.2 Kcal; ±0.03) and were found to adopt allowed conformations (data not shown). The conformation of the 27th frame having minimum most energy was assigned to this region.

The MD simulations of Loop4 of JEVS revealed that the individual sampled conformers are similar in conformation as the rmsd (calculated using the Cα atoms) between them was found to be < 0.37Å. In fact, most of the conformers deviated only by 0.18 Å. This was also seen from the (ϕ,ψ) plot, as these angles varied only occasionally during MD simulation for 500ps (data not shown). The conformation of 46th frame was assigned to Loop4 of JEVS. In the similar fashion, the MD trajectories of Loop5 (365-370) and Loop6 (387-390) were analysed. The sampled conformers of Loop5 were found to deviate around the residues 366, 367, 368. The conformation of the 29th frame (-29.59 Kcal) was assigned to Loop5 and the conformation of the 36th frame (-13.82 Kcal) from the MD trajectory of Loop6 was assigned to Loop6 of JEVS.

3.4 Model of solvated protein

The model of Egp of JEVS was optimised using the distance dependent dielectric constant ($\varepsilon=4r_{ij}$) to simulate the effect of the solvent. The predicted structure was then optimised using the explicit solvent, a 10 Å layer of water. The structure of Egp of JEVS showed little changes upon solvation and the rmsd between the unsolvated and solvated structures of JEVS was found to be 1.26Å for all atoms including side-chains and intra-molecular potential energy has been reduced by about 6%. The side chains were found to be oriented properly and the conformation of the protein was stabilised further by the formation of 957 hydrogen bonds between the protein and water molecules. Thus, 3D structure prediction with explicit water helps to optimise the structure and could be used to understand biological function.

3.5 Evaluation of the model

The 3D model of Egp of JEVS was found to be acceptable geometrically and stereochemically. The PROCHECK G factor for dihedral angles is –0.09 and for main-chain covalent forces is 0.3. The overall average G factor is 0.08 (Laskowski, et. al., 1993). The WHAT_CHECK score of merit is –1.46 (Vriend, 1990).

The combined (pair + surface) energy graphs drawn using ProsaII, (Sippl, 1993) are shown in the Figure 6. The graphs are smoothened by using a window size of 60 residues. As can be seen, the predicted structure of JEVS (red) is threaded properly and overall structure is satisfactory. None of the amino acid residues of Egp of JEVS have positive E/kT value for combined energy and also points out that the structures of Egps of JEVS and JEVN (black) has similar energetic profile, although have been modeled independently.

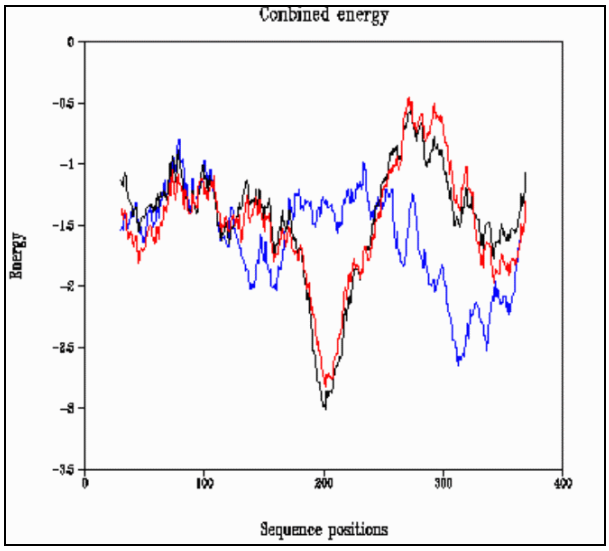

Figure 6: The energy profiles of Egps of TBEV (blue), JEVN (black) and JEVS (red) calculated using ProsaII (Sippl, 1993).

The energy profile of the template structure, Egp of TBEV is shown in blue. As can be seen from Figure 6, the energy of the predicted structures of Egps of JEVS and JEVN is lower than that of the template in the region around residue 200. The predicted structures of Egp of JEVS and JEVN were found to have higher energy for the residues 275-315 as compared to the energy profile of the corresponding region of Egp of TBEV. It could be due to the fact that this region encompasses the domain boundaries. However, even though the energy of this region is higher than the template, it is negative. The ProsaII Z-score for Egp of JEVS was found to be -10.48, which is in the range typical for native folds. The (ϕ,ψ) occupancy in the Ramachandran plot is shown in Figure 7 shown below (Ramachandran & Sasisekharan, 1968).

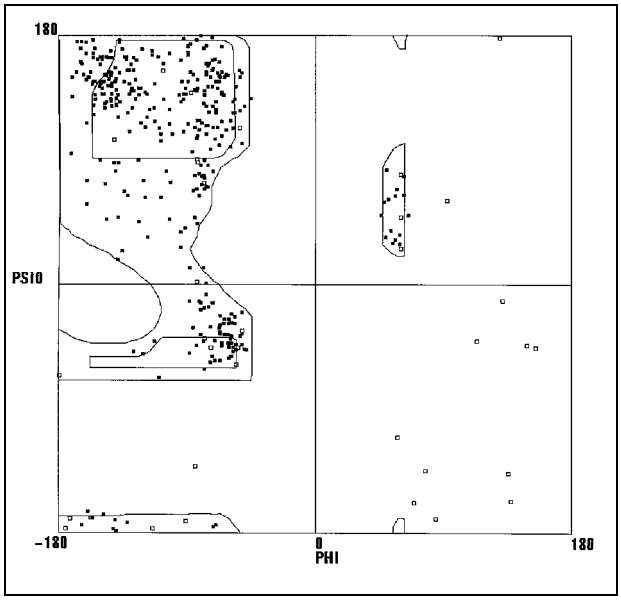

Figure 7: The main chain dihedral angles (ϕ,ψ) are plotted for the predicted 3D structure of Egp of JEVS. The (ϕ,ψ) values of every residue lie within the allowed regions of the Ramachandran plot. ■: Non-Gly amino acids. □: Gly residues.

Thus, the predicted structure of Egp of JEVS was found to be acceptable when evaluated using independent structure evaluation methods.

3.6 Structural description of the model of JEVS

The predicted structure of Egp of JEVS is shown in Figure 8. Superimposition of the predicted structure of Egp of JEVS on the template structure of Egp of TBEV shows very high similarity at 3D structure level between these two proteins, though sequence similarity is only 40%. The rms deviation between the superimposed structure of the template and the model for the Cα trace (390 atoms) and the backbone (1170 atoms) was found to be 1.79Å and 1.98Å respectively. Similar to Egp of TBEV and JEVN, JEVS has an extended structure.

As can be seen from Figure 8, the Egp of JEVS is highly extended and consists of 3 domains. Domain I is referred as the central domain and consists of 128 residues (alignment positions: 1-51; 137-196; 293-311). The second domain, Domain II (also called dimerisation domain, consists of 171 residues (alignment positions: 52-136 and 197-292). Domain III is a contiguous stretch of 100 residues (312-411). There are 30 strands that make 9 sheets, out of which 8 are anti-parallel and one sheet has a mixed topology. JEVS has 2 small additional sheets, one at the base of the dimerisation domain (2 strands: 58-59 & 225-226) and the other in the domain III (2 strands: 335-336, 360-361). There are six additional short strands in JEVS but one strand, which exists in JEVN, was found to be missing in JEVS. It was noted that number of residues in the strands are different in a few cases although overall start and end positions match. The strands were also found to slide and shift by a few residues. As in case of JEVN and TBEV, JEVS has only two alpha helices at positions 215-220 & 256-267 and five 3_{10} helices. JEVS was found to have two additional 3_{10} helices at positions 2-6 and 172-176. However, a 3_{10} helix at position 291-294 in JEVN was found to be missing in JEVS. The backbone and side chain hydrogen bonds were also found to satisfy the geometric criteria.

3.7 Comparison of structures of Egps of JEVS & JEVN

The over all rms deviation between Egp of JEVN and JEVS is 0.84 Å. As has been mentioned, the structures of Egp of JEVN and JEVS were found to deviate by 1.8 and 2.0Å from the template structure of Egp of TBEV where as they deviate only by 0.84Å from each other. The deviations in the structures of Egp of two strains of JEV could be due the effect of the mutations or it could be due to the steps involved in the refinement of the model. Both the structures were found to deviate in the loop regions and the regions that surround mutations.

Loop#: Residues	rmsd (Å)
L1:81-90	0.55
L2:157-161	0.42
L3:272-278	0.99
L4:345-350	0.47
L5:365-370	0.71
L6:387-390	0.67

The rms deviations between the loops of Egp of JEVN and JEVS are listed in the Table 1, which indicate local structural dissimilarity between two structures. At the backbone level, the differences are in the orientations of the loop regions and the secondary structure of loops.

Table 1: rmsd between loops of JEVS & JEVN

The loops were found to adopt various conformations like β-turns, β-hairpins and 3_{10} helices in both, JEVN as well as JEVS. For example, the region 82-86 in JEVN is a 3_{10} helix, which is followed by the type IV β-turn at position 87-90.

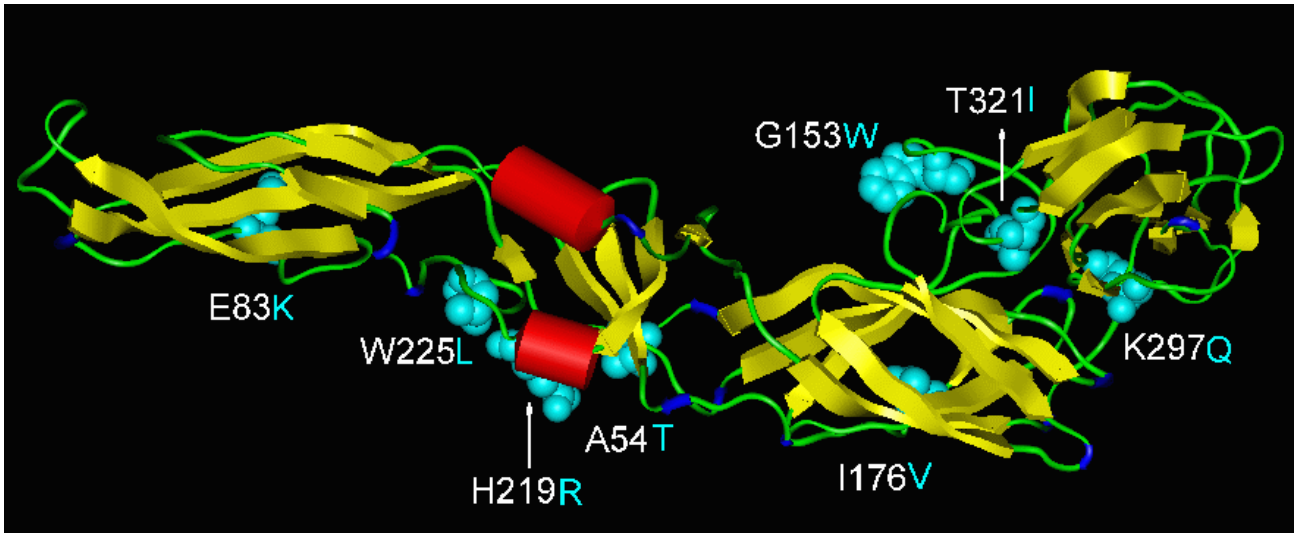

Figure 8: Predicted structure of Egp of JEVS. The β-sheets form the major secondary structure. Two short α-helices are present. The strands, helices, turns and coils are shown in yellow, red, blue and green respectively. The mutant residues are shown in CPK rendering (cyan). The amino acids and residue numbers in JEVN are shown in white and mutations in JEVS are shown in cyan.

The position of 3_{10} helix in JEVS has been shifted by a residue and the following region could not be assigned any regular secondary structure. Furthermore, the structure of Loop1 in JEVS was found to be energetically more stable than the Loop1 in JEVN (data not shown). The Glu83 in JEVN has been replaced by Lys in JEVS. Both these residues are charged and have opposite charges (Glu: negatively charged and Lys: positively charged). The effect of this mutation of the overall structure is discussed in the following section. The region 365-369, which has 100% sequence identity in both strains of JEV, was found to adopt type I β-turn in JEVN and is type IV β-turn in JEVS respectively. Thus, the MD simulations approach used to predict the conformations of the loop regions seems to be a proper method to study the local structural movements in the related proteins. The regions, 157-160, 274-277 in JEVN as well as JEVS was found to adopt type I and type IV β-turn respectively. Similarly, the region 388-391 was found to adopt a hairpin conformation in JEVN as well as in JEVS.

When the rmsd of the backbone atoms between the JEVN and JEVS was calculated, it was found that the mutation, G153W brings the change in the backbone structure of the region, which is 5 residues upstream and downstream of 153rd position by about 1.09Å. Similarly, the mutation A54T, which is a part of SCR region, has more effect on the structure of the 5 residues downstream (1.01 Å) than on the 5 residues upstream (0.6 Å). In case of the mutation K297Q the maximum deviation was observed around the site of mutation and 5 residues upstream contribute maximally to the rmsd value.

The effect of the mutations on the structures of Egp of JEVN and JEVS was evident at the tertiary structure level as well. The accessibility (%) at two positions E83K and H219R were found to change by more than 10% (Lee & Richards, 1971). Such residues could be playing an important role in the recognition of these strains by the immune system and could be one of the major determinants responsible to bring out the strain specificity.

Similarly, the sequence and structural information around the sites of mutations and in the loop regions, if transferred into the profile of chemical properties, could lead to a few interesting observations, which could further explain the virus specific properties. For example, the sequence of Loop1 in TBEV is TLAEEHQGG, which has two negatively charged residues (EE) that are followed by H, a polar residue, which gets positively charged. The sequence of the same region in the JEVN (HNEKRADSS) has 3 positively charged residues (H, K, R) and the pattern (EEH: 2 negative + 1 positive charges) in TBEV is changed to (EKR: 1 negative + 2 positive charges) in JEVN. This would imply that, if this is the region that is recognized by the immune system, then the antibodies developed against TBEV and JEVN would require exactly opposite pattern of charges in their CDR regions. Furthermore, the same loop in the JEVS has a mutation at position 83, where E in JEVN is replaced by K (HNKKRADSS). This mutation removes the only negative charge in the JEVN and makes it entirely positively charged region in JEVS and thus becomes strain specific.

Similarly the region (149-164) encompasses the Loop2 and is one of the hyper variable regions amongst flaviviruses (multiple alignment data not shown). The mutation G153W in JEVN and JEVS is part of this region. The TBEV has Thr at the equivalent position, which is a small amino acid. It is substituted by the Gly residue, which is the smallest amino acid having hydrogen atom as a side chain, in JEVN. Thus, the small amino acid residues in TBEV and JEVN are substituted by aromatic residue W in JEVS, which has hydrophobic character and is bulky in size. As has been noted in the Table 1, the backbone deviates from its position to accommodate W.

The sites of mutations in the predicted structures of Egp of JEVN and JEVS were also analysed to study various atomic interactions of the mutated residues and the surrounding residues (Sobolev, et al., 1999).

Virus: Mutant residue	# IR	Types of interactions			
		HB	AC	HC	DC
JEVN:E83	5	4	-	-	4
JEVS:K83	5	4	-	7	4
TBEV:A83	4	5	-	-	4

Table 2: Interatomic contacts of mutant residues calculated using (Sobolev, et al., 1999)

The interactions involving the equivalent residues from the Egp of TBEV were also compiled. The number of interacting residues (IR) and types of interactions like hydrogen bonds (HB), aromatic contact (AC), hydrophobic contact (HC) and destabilizing contact (DC) have been compiled for every mutant. As an example, the contact for the mutation E83K have been listed in Table 2. The negatively charged E83 in JEVN has been substituted by K in JEVS, which is a positively charged polar residue. The K83 in JEVS was found to have 7 HC contacts, 5 with K84 and 2 with A86. The TBEV has Ala at the structurally equivalent position and it was found to interact with the neighbouring residues only by forming hydrogen bonds (HB).

Thus, the predicted structures of Egps of two strains of JEV were analysed to study the strain specific properties of these viruses. The B-cell epitopes were predicted using Kolaskar & Tongaonkar (1990) method. The conformational and sequential antigenic determinants of Egp of JEV were predicted using approach developed (Kulkarni-Kale, 2002, Kolaskar & Kulkarni-Kale, 1999) and the 3D structure data of Egp of JEVS. The antigenic determinants predicted using this approach are shown in the Figure 1. A few unique strain specific determinants like (NTEAF) and (TSSANSK) were predicted for JEVS. The predicted antigenic determinants have been used to identify a candidate peptide to design synthetic peptide vaccine against JEV. A strategy to design peptide vaccines is also being developed, which involves prediction of the conformations of antigenic peptides

using 'Cut & Stitch' approach, molecular dynamics simulations of nanosecond durations using an approach developed by Kolaskar & Sawant (unpublished) and identification of a polypeptide spacer to join B and T cell epitopes (Kolaskar & Kulkarni-Kale, unpublished). The peptides have been synthesized in the laboratory by our collaborators and the experimental results show that these peptides do induce the neutralizing antibodies, which are capable of neutralising the virus, when challenged (Dewasthaly, Gore, Kulkarni-Kale & Kolaskar, unpublished data), which clearly demonstrates how *in-silico* simulations facilitate wet lab experiments.

4 Acknowledgements

This work was supported by the grant from Department of Biotechnology, Government of India.

5 References

BAKER, D and SALI, A. (2001): Protein structure prediction and structural genomics. Science. 294:93-96.

BANERJEE, K. (1996): Emerging viral infections with special reference to India. Indian Journal of Medical Research. 103:177-200.

BERENDSEN, H.J.C. and HAYWARD, S. (2000): Collective protein dynamics in relation to function. Curr. Op. Struc. Biol. 10:165-169.

BURLEY, S.K. and BONANNO, J.B. (2002): Structuring the universe of proteins. Annu Rev Genomics Hum Genet. 3:243-362.

CHAMBERS, T.J., TSAI, T.F., PERVIKOV, Y., MONATH, T.P. (1997): Vaccine development against dengue and Japanese encephalitis: report of a World Health Organization meeting. Vaccine. 15:1494-1502.

DALEY, A.J. and DWYER, D.E. (2002): Emerging viral infections in Australia. J Paediatr Child Health. 38:1-3.

DONIACH, S. and EASTMAN, P. (1999): protein dynamics simulations from nanoseconds to microseconds. Curr. Op. Struc. Biol. 9:157-163.

DYSON, H.J. and WRIGHT, P.E. (1995): Antigenic peptides. FASEB J. 9:37-42.

EASTON, A. (1999): Outbreak of Japanese encephalitis hits Malaysia. BMJ. 318:893.

ELLIS, P.M., DANIELS, P.W., BANKS, D.J. (2000): Japanese encephalitis. Vet Clin North Am Equine Pract.16:565-578.

GAMBEL, J.M., DEFRAITES, R., HOKE, C.H., BROWN, A., SANCHEZ, J., KARABATSOS, N., TSAI, T., MESCHIEVITZ, C. (1995): Japanese encephalitis vaccine: persistence of antibody up to 3 years after a three-dose primary series. J Infect Dis. 171:1074.

GOWAL, D., SINGH, G., BHAU, L.N, SAXENA, S.N. (1991). Thermostability of Japanese encephalitis vaccine produced in India. Biologicals. 19:37-40.

HOKE, C.H., NISALAK, A., SANGAWHIPA, N., JATANASEN, S., LAORAKAPONGSE, T., INNIS, B.L., KOTCHASENEE, S., GINGRICH, J.B., LATENDRESSE, J., FUKAI, K., et al. (1988): Protection against Japanese encephalitis by inactivated vaccines. N. Engl. J. Med. 319: 608-614.

JOHNSON, M.S., SRINIVASAN, N., SOWDHAMINI, R., BLUNDELL, T.L. (1994): Knowledge-based protein modeling. Crit. Rev. Biochem. Mol. Biol. 29**:**1-68.

KOLASKAR, A.S. and KULKARNI-KALE, U. (1992): Sequence alignment approach to pick up conformationally similar protein fragments. J. Mol. Biol. 223:1053-1061.

KOLASKAR, A.S. and 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.

KOLASKAR, A.S. and TONGAONKAR, P.C. (1990): A semi-empirical method for prediction of antigenic determinants on protein antigens. FEBS Lett. 276:172-174.

KULKARNI-KALE, U. (2002). Prediction of structures and functions of proteins. Ph.D. Thesis. University of Pune, Pune.

LASKOWSKI, R.A., MACARTHUR, M.W., MOSS, D. S., THORNTON, J.M. (1993): Procheck: a program to check the stereochemical quality of protein structure. J. Appl. Cryst. 26:283-291.

LEE, B. and RICHARDS, F.M. (1971): The interpretation of protein structures: estimation of static accessibility. J. Mol. Biol. 55:379-400.

LIU, J. and ROST, B. (2002): Target space for structural genomics revisited. Bioinformatics. 18:922-933.

MARTI-RENOM, M.A., MADHUSUDHAN, M.S., FISER, A., ROST, B., SALI, A. (2002): Reliability of assessment of protein structure prediction methods. Structure (Camb). 10:435-440.

MCCORMACK, J.G. and ALLWORTH, A.M. (2002): 8: Emerging viral infections in Australia. Med J Aust. 177:45-49.

MOHAN RAO, C.V., RISBUD, A.R., DANDAWATE, C.N., UMARANI, U.B., AYACHIT, V.M. RODRIGUES, F.M., PAVRI, K. M. (1993): Serological response to Japanese encephalitis vaccine in a group of school children in South Arcot district of Tamil Nadu. Indian Journal of Medical Research 97: 53-59.

MONATH, T.P. (2002): Japanese encephalitis vaccines: current vaccines and future prospects. Curr Top Microbiol Immunol.267:105-138.

MONATH, T.P. and HEINZ, F.X. (1996): Flaviviruses. In Fields Virology. 961-1043. (B. N. Fields, D. M. Knipe, P.M. Howley et.al., Eds), Lippincott-Raven Publishers, Philadelphia.

MOULT, J., HUBBARD, T., BRYANT, S.H., FIDELIS, K., PEDERSEN, J.T. (1997): Critical assessment of methods of protein structure prediction (CASP): round II. Proteins. Suppl 1:2-6.

OFFRINGA, R., VAN DER BURG, S.H., OSSENDORP, F., TOES, R.E., MELIEF, C.J. (2000): Design and evaluation of antigen-specific vaccination strategies against cancer. Curr Opin Immunol. 12:576-582.

PARANJPE, S., and BANERJEE, K. (1996): Phylogenetic analysis of the envelope gene of Japanese encephalitis virus. Virus Res. 42:107-117.

PEITSCH, M.C. (2002): About the use of protein models. Bioinformatics. 18:934-938.

RAMACHANDRAN, G.N. and SASISEKHARAN, V., (1968). Conformation of polypeptides and proteins., Adv. Protein Chem. 23:283-438.

REY, F.A., HEINZ, F.X., MANDL, C., KUNZ, C., HARRISON, S.C. (1995): The envelope glycoprotein from tick-borne encephalitis virus at 2 A resolution. Nature 375: 291-298.

SHLIM, D.R. and SOLOMON T. (2002): Japanese encephalitis vaccine for travelers: exploring the limits of risk. Clin Infect Dis. 35:183-188.

SIPPL, M.J. (1993): Recognition of errors in three-dimensional structures of proteins. Proteins 17, 355-362.

SOBOLEV, V., SOROKINE, A., PRILUSKY, J., ABOLA, E.E., EDELMAN, M. (1999): Automated analysis of interatomic contacts in proteins. Bioinformatics. 15:327-332.

TSAI, T.F. (2000): New initiatives for the control of Japanese encephalitis by vaccination: minutes of a WHO/CVI meeting, Bangkok, Thailand, 13-15 October 1998. Vaccine. 18:1-25.

VAN DEN HURK, A.F., NISBET, D.J., JOHANSEN, C.A., FOLEY, P.N., RITCHIE, S.A., MACKENZIE, J.S. (2001): Japanese encephalitis on Badu Island, Australia: the first isolation of Japanese encephalitis virus from Culex gelidus in the Australasian region and the role of mosquito host-feeding patterns in virus transmission cycles. Trans R Soc Trop Med Hyg. 95:595-600.

VAN REGENMORTEL, M.H.V., FAUQUET, C.M., BISHOP, D.H.L., CARSTENS, E.B., ESTES, M.K., LEMON, S.M., MANILOFF, J. (2000): Virus Taxonomy: The classification and nomenclature of viruses. The 7th report of the International Committee on Taxonomy of Viruses.

VENCLOVAS, C., ZEMLA, A., FIDELIS, K., MOULT, J. (2001): Comparison of performance in successive CASP experiments. Proteins. Suppl 5:163-170.

VENUGOPAL, K. and GOULD, E.A. (1994): Towards a new generation of flavivirus vaccines. Vaccine. 12:966-975.

VRIEND, G. (1990): WHAT IF: a molecular modeling and drug design program. J. Mol. Graph. 8, 52-56.

WEBER, J. (2002): Peptide vaccines for cancer. Cancer Invest. 20:208-221.

WISDOM, G.B.(1992):Peptide antigens. Biochem Soc Trans. 20:226-229.