African Journal of Biotechnology Vol. 10(18), pp. 3528-3533, 2 May, 2011
Available online at http://www.academicjournals.org/AJB
DOI: 10.5897/AJB10.132
ISSN 1684–5315 © 2011 Academic Journals

Full Length Research Paper

Bioinformatics Analysis of Envelope Glycoprotein E epitopes of Dengue Virus Type 3

Muhammad Ilyas[1], Ziaur Rahman[2], Sulaiman Shamas[1], Mukhtar Alam[3], Muhammad Israr[1] and Khalid Masood[1*]

[1]National Centre of Excellence in Molecular Biology, University of the Punjab Lahore, Pakistan.
[2]Department of Biological Sciences, Korea Advanced Institute of Science and Technology, Daejeon, 305-701. South Korea.
[3]Hazara University, Mansehra, Khyber Pakhtoonkhwa, Pakistan.

Accepted 17 December, 2010

The E glycoprotein of dengue virus is responsible for the viral binding to the receptor. The crystal structure of envelope glycoprotein has already been determined. However, where the well-defined B-cell and T-cell epitopes are located is still a question. Because of the large variations among the four dengue genotypes, it is very hard to design conserved epitopes for all of them. Therefore, we selected only one genotype (DENV3). The conserved regions were found in more than 600 DENV E glycoprotein sequences. Both the B-cell and T-cell epitopes were predicted and the hydrophobicity, antigenicity, accessibility and flexibility of the highly conserved E glycoprotein were further predicted by using different bioinformatics algorithms. The secondary structure was obtained and the predicted epitopes were pointed out in it. Binding sites on glycoprotein of DENV-3 for attachment of virus to the receptor was identified, while keeping those attachments in which new drugs for dengue related infections could not be designed.

Key words: Glycoprotein, DENV3, epitopes, antigenecity, B-cell, T-cell, vaccine.

INTRODUCTION

Approximately, 3 billion people are at risk of acquiring dengue viral infection in more than 100 countries in tropical and subtropical regions. Annually, it is estimated that 100 million cases of dengue fever (DF) and half a million cases of dengue haemorrhagic fever/dengue shock syndrome (DHF/DSS) occur worldwide resulting in approximately 25,000 deaths (WHO, 2002). Dengue is caused by any of four related viruses (DENV-1, -2, -3 and -4) transmitted by mosquitoes. Dengue virus (DENV) is an ssRNA positive-strand virus of the family Flaviviridae; genus Flavivirus. There are four serotypes of DENV. The virus has a genome of about 11000 bases that code for three structural proteins: C, prM and E; seven non-structural proteins: NS1, NS2a, NS2b, NS3, NS4a, NS4b and NS5; and short non-coding regions on both the 5' and 3' ends (Mackenzie et al., 2004). No prophylactic vaccine is currently available to prevent infection of DENV and the most effective protective measures are those that avoid mosquito bites. There are many ongoing vaccine development programs. Among them is the Pediatric Dengue Vaccine Initiative set up in 2003 with the aim of accelerating the development and introduction of dengue vaccine(s) that are affordable and accessible to poor children in endemic countries (Halstead and Deen, 2002). Increased efforts are therefore needed in the development of an effective vaccine against DENV. The envelope E glycoprotein having 495 amino acids was reported to play an important role in the DENV attachment

*Corresponding author. E-mail: khalid@cemb.edu.pk.

Abbreviations: DHF/DSS, Dengue haemorrhagic fever/dengue shock syndrome; **DF,** dengue fever; **DENV,** dengue virus; **NCBI,** national center for biotechnology information; **BLAST,** basic local alignment search tool; **VAST,** vector alignment search tool.

Table 1. Predicted B-cell epitopes of DENV-3 glycoprotein E.

S/N	Start Position	Sequence	End Position
1	17	GATWVDVVLEHGGCVT	32
2	50	ATQLATLRKLCIE	62
3	87	DQNYVCKHTY	96
4	210	WFFDLPLPW	218
5	232	KELLVTF	238
6	244	KKQEVVVLG	252
7	315	HGTILIKV	322
8	350	ITANPVVT	357
9	422	SVGGVLN	428

to the host cell receptors and entry into the target cells (Roehrig, 1997). Therefore, it is one of the most valuable candidate proteins for the development of DENV vaccine. With the development of bioinformatics, we havesome tools and methods to study the rules in the variation of E glycoprotein based on the fact that as an important structural protein, there must be some conservative regions to maintain its function and structure, and by the analysis of huge number of sequences of envelop protein, we may find some rules in the variation and then can predict the future sequence and structure to design new vaccine for the prevention of DENV infection. Parameters such as hydrophobicity, flexibility, accessibility, antigenicity and exposed surface of polypeptide chains have been correlated with the location of continuous epitopes in a few well characterized proteins. Web servers and other softwares were used in our study.

MATERIALS AND METHODS

Searching for serotype sequences of DENV-3

We searched the National Center for Biotechnology Information (NCBI) website for the E glycoprotein of Dengue serotype 3. Total number of sequences found was 1217, in which more than 600 were randomly selected. The sequences format utilized for saving downloaded E protein sequences in each DENV serotypes was FASTA format. This format was submitted for query of further analysis. All the selected sequences were compared to find mutational and conservative regions by using Clustal W software (Thompson et al., 1994).

B-cell epitopes of E DENV-3 protein prediction

B cell epitopes prediction was done by using Antigen Prediction server, which can be accessed freely in their website at http://bio.dfci.harvard.edu/Tools. Antigenic peptides are determined using the method of Kolaskar and Tongaonkar (1990). Antigen Prediction server needs protein sequence data as its input for B cell epitope prediction. The prediction result was peptide sequences with their start and end position in the E DENV-3 protein sequences.

T-cell epitopes of E DENV-3 protein prediction

T cell epitopes prediction was done with a neural network based MHC Class-I Binding Peptide Prediction Server (nHLAPred) (Bhasin and Raghava, 2006). The server website is http://www.imtech.res.in/raghava/nhlapred/. nHLAPred needs data input of E DENV-3 protein sequences in order to predict the epitope. The epitope determination was done based on amino acid sequences selection which have the highest binding score.

Secondary structure prediction

Secondary structure of E DENV-3 was predicted using SOPMA library (Geourjon and Deléage, 1995), which can be accessed by going to their freely available server (http://npsapbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_hnn.html).

Hydrophobicity, accessibility, flexibility and antigenecity of E DENV-3

Protein package in DNASTAR software (Burland, 2000) helps to predict and display different patterns, like secondary structural characteristics and physicochemical properties of protein sequences via its comprehensive suite of protein analysis tools. Protein's simple, elegant and graphical user interface can also help us locate antigenic determinants and predict protease digestion patterns.

Template inquiry and tertiary structure homology modeling

The complete x-ray crystallographic structure of the Dengue virus type-3 envelope glycoprotein is known and well characterized (Modis et al., 2005). The structure was downloaded from Protein Data Bank (PDB ID: 1UZG.pdb) and was further studied with the help of SWISSPDB viewer (Guex and Peitsch, 1997) and Chimera (Pettersen et al., 2004).

RESULTS

B-cell and T-cell epitopes prediction

The predicted results for both antigenic peptides of the

Table 2. Predicted T-cell epitopes of DENV-3 glycoprotein E.

HLA sites		Peptides position	
HLA-A2	106-114	171-179	291-299
HLA-A11	30-38	237-245	274-282
HLA-A24	298-306	388-396	479-487
HLA-B51	204-212	210-218	389-397
HLA-B60	48-56	181-189	254-262
HLA-B62	412-420	45-53	268-276

envelope glycoprotein for DENV-3 are shown in Tables 1 and 2. The B-cell (Table 1) and the T-cell epitopes (Table 2) of the glycoprotein of DENV-3 were predicted by using freely accessible online servers.

Protein sequence analysis

The results obtained from the protein package in DNASTAR software is shown in Figure 2. Algorithms used in this analysis were hydrophobicity plot (Kyte and Doolittle, 1982), flexible regions (Karplus and Schultz, 1985), antigenic index (Jameson and Wolf, 1988) and surface accessibility (Emini et al., 1985).

Structure prediction

Secondary structure was predicted using SOPMA server shown in Figure 1. In the result, it was observed that 20.89% regions in the structure was alpha helix, 6.09% was beta turns and 40.57% was coils. The result was approximately similar in all the sequences. As there was no big difference in the sequences, it was thought that the conserved regions maintained the special structure of E glycoprotein of DENV-3. The tertiary structure was retrieved from the PDB server (PDB ID: 1UZG.pdb) (Figures 3A and B).

DISCUSSION

As a result of the great mutation sites, it is very hard to design specific drug for dengue related infections (Aquino et al., 2008). However, many researchers are trying to find out a way to develop drug for treatment and vaccination of dengue infections (Anthony et al., 2009). E DENV-3 proteins were chosen because of their high prevalence and death-cause of many individuals in the Pakistani region (Jamil et al., 2007). This is the reason that we focused on type 3 of dengue virus. A better drug can be designed if individual serotype is studied.

In the current study, we analyzed dengue virus variant region of polyprotein, that is, envelope protein. The nature More than 600 envelop glycoprotein sequences were of this protein was analyzed with computational software .aligned, some variable regions were found. It was observed that 75.2% of the bases were conserved and the mutation rate was approximately 20%. A large number of conservative sites were found in the glycoprotein. Great number of polar amino acids (Ser, Thr, Asn and Gln) were observed, while glycine was reported as the most available amino acid in the conserved regions.

On the basis of prediction of secondary structure, hydrophilicity, accessibility, antigenicity and flexibility methods were selected to predict the epitopes. We have also found some highly conserved antigenic epitopes present in the protein, which may be useful for the diagnosis of different variants of dengue. Both B-cell and T-cell epitopes of dengue virus type 3 envelop E glycoprotein were developed by bioinformatics approaches, the sequences at 37-66, 80-121, 183-228 and 242-279 amino acid (aa) for dengue virus type 3 were predicted as the more prevalent epitopes by using multiple parameters and different analysis softwares, respectively. The amino acid sequences with high binding score have a high possibility of inducing antibody response (Tambunan et al., 2009). In this study, the sequences selected not only have higher scores in the average antigen index (AI), which could predict the antigen epitope of envelop glycoprotein E, but also showed better hydrophilic properties. The Basic Local Alignment Search Tool (BLAST) and Vector Alignment Search Tool (VAST) score results of the vaccines by Tambunan et al. (2009) have shown more than 90% identity with E DENV2 and DENV3 protein.

In the protein secondary structure prediction, it was seen that there are β-turn and alpha-helix regions, which are stable structure beside conservative sites, such as 53-63, 116-121, 256-267, 326-334, 425-438 and 450-459. So it was thought that, it is these regions that maintain the stable structures of envelop protein. Despite the fact that the DENV envelop is the most variable region of the genome, some conservative sites were still found and the possible B-cell and T-cell epitopes were predicted. The peptides of E-protein explained herein compared with other well documented epitopes (Rangel et al., 2009) are potentially relevant for the development of an effective vaccine for the dengue virus.

Conclusion

In this research, some epitopes in the structural protein of dengue virus type 3 variants was predicted. The conserved epitopes may be useful for the diagnosis of the dengue type 3 variant. It is believed that all of these findings will prompt the development of dengue vaccine and more research should be focused on the structure of E glycoprotein and its interaction with antibody with the method of molecular dynamics and molecular model

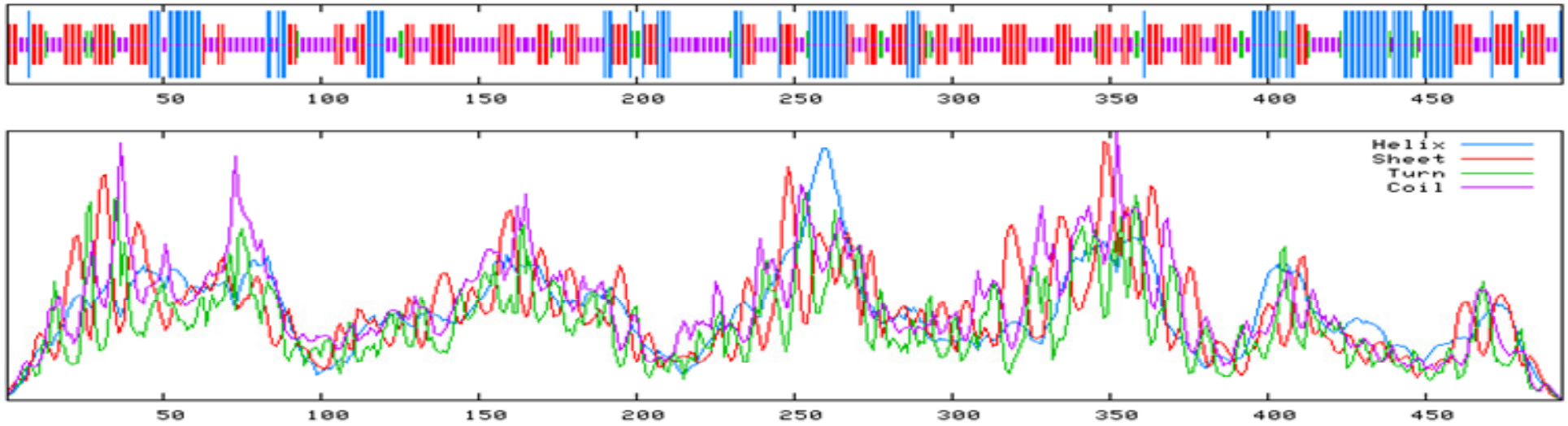

Figure 1. Secondary structure of DENV-3 glycoprotein E.

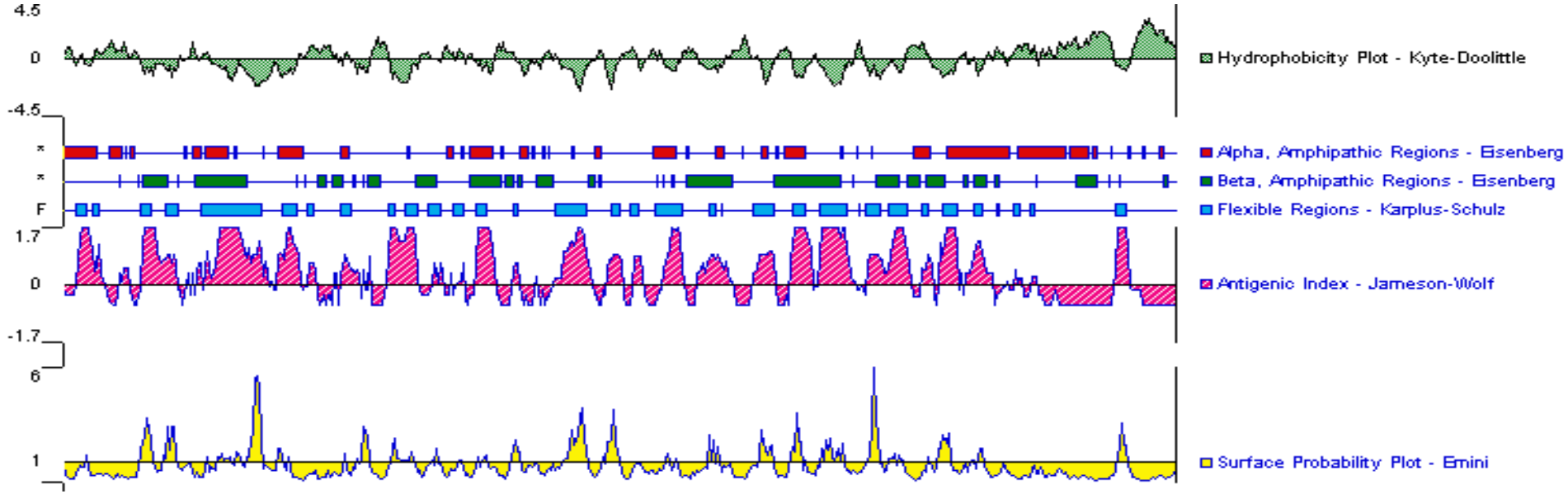

Figure 2. Hydrophobicity, flexibility, antigenicity and surface probability of the DENV-3 glycoprotein E.

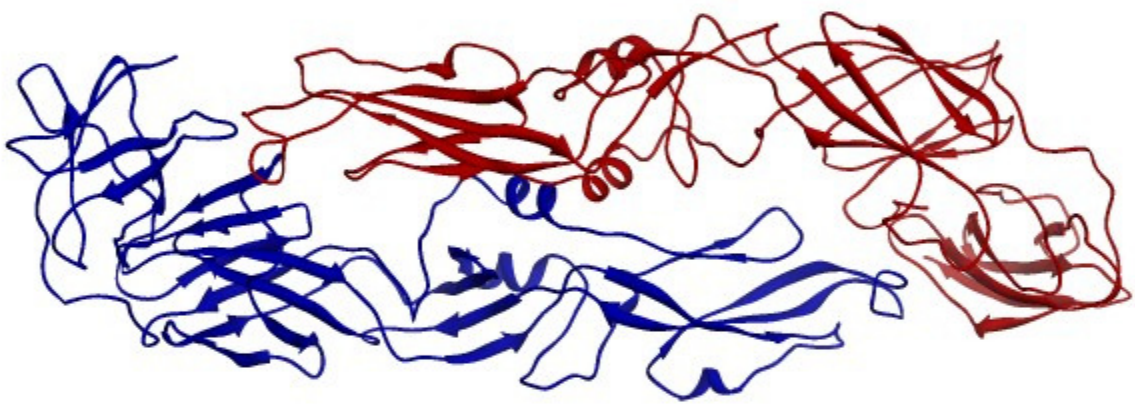

Figure 3 (A): Tertiary structure of the DENV-3 glycoprotein E.

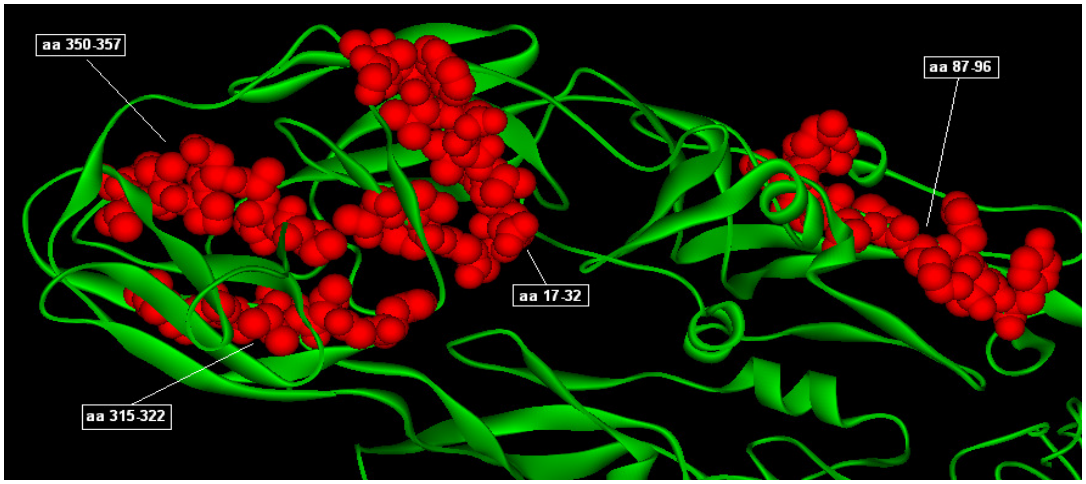

Figure 3 (B): Visualizing epitopes positions in Tertiary structure of the DENV-3 glycoprotein E.

The predicted epitopes may be used for vaccine development against dengue hemorrhagic fever.

ACKNOWLEDGEMENTS

We would like to express our thanks to Dr. Jong Bhak of Korean Bio Information Center (KOBIC) who encouraged us by showing interest in our work. He generously shared his knowledge with us. We are thankful to Korean Bio Information Center (KOBIC) for facilitating us during this work.

REFERENCES

Anthony KG, Bai F, Krishnan MN, Fikrig E, Koski RA (2009). Effective siRNA targeting of the 3' untranslated region of the West Nile virus genome. Antiviral Res. 82: 166-168

Aquino JD, Tang WF, Ishii R, Ono T, Eshita Y, Aono H, Makino Y (2008). Molecular epidemiology of dengue virus serotypes 2 and 3 in Paraguay during 2001-2006: The association of viral clade introductions with shifting serotype dominance. Virus Res. 137(2): 266-270

Bhasin M, Raghava GPS (2006). A hybrid approach for predicting promiscuous MHC class I restricted T cell epitopes; J. Biosci. 32: 31-42

Burland TG (2000): DNASTAR's Lasergene sequence analysis software. Methods Mol. Biol. 132: 71-91

Emini EA, Hughes JV, Perlow DS, Boger J (1985). Induction of hepatitis A virus-neutralizing antibody by a virus-specific synthetic peptide. J. Virol. 55(3): 836-839.

Geourjon C, Deléage G (1995). SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. Bioinformatics, 11(6): 681-684

Guex N, Peitsch MC (1997). SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. Electrophoresis, 18(15): 2714-2723.

Halstead SB, Deen J (2002). The future of dengue vaccines. Lancet, 360: 1243-1245

Jameson BA, Wolf H (1988). The Antigenic Index, A novel algorithm for predicting antigenic determinants. Comput. Appl. Biosci. 4: 181-186.

Jamil B, Hasan R, Zafar A, Bewley K, Chamberlain J, Mioulet V, Rowlands M, Hewson R (2007). Endemic Dengue Fever: a seldom recognized hazard for Pakistani children. Emerg Infect Dis. 13: 182-183.

Karplus PA, Schultz GE (1985). Prediction of chain flexibility of peptide

antigens, Naturwissenchaften, 72: 212-213.
Kolaskar AS, Tongaonkar PC (1990). A semi-empirical method for prediction of antigenic determinants on protein antigens, FEBS Lett. 276: 172-174
Kyte J, Doolittle RF (1982). Hydropathies listed above were taken from the Hydropathy scale derived. Jol. Mol. Bio. 157: 105-132.
Mackenzie J, Gubler D, Petersen L (2004). Emerging flaviviruses: the spread and resurgence of Japanese encephalitis, West Nile and dengue viruses. Nat. Med. 10: 98-109.
Modis Y, Ogata S, Clements D, Harrison SC (2005). Variable surface epitopes in the crystal structure of dengue virus type 3 envelope glycoprotein. J. Virol. 79(2): 1223-1231.
Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE (2004) UCSF Chimera-a visualization system for exploratory research and analysis. J. Comput. Chem. 25(13): 1605-1612.
Roehrig JT (1997). Immunochemistry of the dengue viruses. In: Gubler D J, Kuno G, editors. Dengue and dengue hemorrhagic fever. New York, N.Y: CAB International. pp. 199-219.
.
Rangel da Silva ANM, Nascimento EJM, Cordeiro MT, Gil LHVG, Abath FGC, Montenegro SM, Marques ET (2009). Identification of Continuous Human B-Cell Epitopes in the Envelope Glycoprotein of Dengue Virus Type 3 (DENV-3). PLoS ONE, 4(10): 7425Thompson JD, Higgins DG, Gibson TJ (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22: 4673-4680.
Tambunan USF, Parikesit AA, Hendra, Taufik RI, Amelia F, Syamsudin (2009). In Silico Analysis of Envelope Dengue Virus-2 and Envelope Dengue Virus-3 Protein as the Backbone of Dengue Virus Tetravalent Vaccine by Using Homology Modeling Method. OnLine J. Biol. Sci. 9(1): 6-16
WHO (2002). World Health Organization. Dengue and Dengue Haemorrhagic Fever. Fact Sheet. Geneva. p. 117.