

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

Usman Sumo Friend Tambunan, Arli Aditya Parikesit, Hendra,
Rizky Ichsan Taufik, Fitri Amelia and Syamsudin
Department of Chemistry, Faculty of Mathematics and Natural Science,
University of Indonesia, Depok 16424, Indonesia

Abstract: Problem statement: Dengue fever, which was caused by Dengue virus infection, had become a major public health problem in the tropic and subtropical countries. Dengue virus (DENV) had four serotypes (DENV-1, DENV-2, DENV-3 and DENV-4), based on their immunogenic in the human body. Preventive measure will be necessary to decrease the prevalence of dengue fever, by developing modern vaccine. **Approach:** This research was focused on in silico study of dengue virus vaccines, by using envelope (E) protein of DENV-2 and DENV-3 as their backbones. T cell epitope prediction was determined by using MULTIPRED server and B cell epitope prediction was determined by using Conformational Epitope Prediction server (CEP). Homology modeling study of E DENV-3 protein as the vaccine backbone had produced six dengue vaccine peptides (HMM Vaccine 1-6). Moreover, homology modeling study of E DENV-2 protein as vaccine backbone had produced six dengue vaccine peptides (ANN vaccine 1-6). **Results:** The BLAST analysis of HMM and ANN vaccines had produced 93% and 91% identity, respectively. The Ramachandran Plot of both vaccines had shown less than 15% non glycine residue in the disallowed region, therefore it showed the solid stability of the proteins. The VAST analysis of E DENV-3 backbone vaccines had determined, that HMM4 and HMM6 had the highest structure similarity with native E DENV-3. HMM4 and HMM6 had the highest VAST score of 64.5. Moreover, the VAST analysis of E DENV-2 backbone vaccines had determined, that ANN1, ANN3, ANN4, ANN5 and ANN6 had the highest structure similarity with native E DENV-2. ANN1, ANN3, ANN4, ANN5 and ANN6 have the highest VAST score of 64.7. **Conclusion/Recommendation:** It could be inferred from this research that HMM4; HMM6; ANN1; ANN3; ANN4; ANN5; and ANN6 were the best in silico vaccine design, based on their similarity with native E DENV Proteins. This research could be applied for the wet laboratory and computerised vaccine design.

Key words: Conformational epitope prediction, dengue fever, dengue vaccine, envelope protein, homology modeling

INTRODUCTION

Dengue fever is an acute febrile viral disease characterized by sudden onset, fever of 3-5 days, intense headache, myalgia, anthralgic retro-orbital pain, anorexia, GI disturbances and rash. Dengue viruses are member of flavivirus family. They included four serotypes of 1, 2, 3 and 4 (Dengue- 1, -2, -3 and - 4). These viruses are also responsible for Dengue Hemorrhagic Fever (DHF). The viruses are transmitted to human by the bite of infective mosquitoes, mainly *Aedes aegypti*. The incubation period is 4-7 days (range 3-14 days). This disease is now endemic in most tropical countries. DHF is characterized by increased

vascular permeability, hypovolaemia and abnormal blood clotting mechanisms^[1].

Dengue Fever (DF), with its severe manifestations such as DHF and Dengue Shock Syndrome (DSS), has emerged as a major public health problem of international concern. The geographical distribution has greatly expanded over the last 30 years, because of increased potential for breeding of *Aedes aegypti*, the vector species. This has been prompted by demographic explosion, rapid growth of urban centers with a strain on public services, such as potable water. Breeding of *Aedes aegypti* was expanding rapidly due to rainwater storing in diverse types of containers^[2].

Corresponding Author: Usman Sumo Friend Tambunan, Department of Chemistry, Faculty of Mathematics and Natural Science, University of Indonesia, Depok 16424, Indonesia

About 100 countries are endemic for DHF and about 40% of the world population (2.5 billion people) is at risk in the tropics and sub-tropics region. Over 50 million DF infections with about 400,000 cases of DHF are reported annually which is a leading cause of childhood mortality in several Asian countries^[2,3].

According to Indonesian Department of Health fact sheet, DF is on the rank 8 of 10 infection diseases, which are considered important in the budget priority and political commitment. All of the dengue virus serotype (DENV) are endemic in many of the big cities and occurred at yearly basis^[4,5].

The vaccine of DF is not yet available. The overwhelming obstacles of the dengue vaccine development are the unknown pathogenesis of dengue virus in human host and the difficulty of virus growth in the culture medium. The animal testing was done for immunogenic and disease prevention purpose only^[1,3,6].

An attenuated vaccine is under clinical trial on human subject now. The vaccine is developed by Mahidol University at Bangkok and Walter Reed Army Institute of Research at the United States. There are risks that could occur by misusing attenuated vaccine. If the virus in the vaccine is not attenuated enough, the vaccine will attack the human host like the DENV infection. Moreover, if the vaccine contains over attenuated virus, the vaccine would not induce the body immune response^[3,7].

The modern engineered vaccine for DENV infection is necessary; in order to overcome the problems of attenuated vaccine. One of the engineered vaccine types is protein/peptide vaccine. Protein/peptide vaccine consists of peptide sequence which derived from the small part of pathogenic protein. There are antigenic determinant in certain part of the vaccine peptide sequences^[8-10].

The multiple alignment study of dengue virus tetravalent vaccine design has been done by our group. The multiple alignments that were done with 102 dengue virus intra serotype of the four DENV serotype concluded high similarity of E protein among each of DENV intra serotype. Although mutation occurred on E protein, the antibody can recognize all DENV intra serotype as one dengue virus serotype^[3,11,12].

The general objective of this research is to design dengue virus vaccines with in silico method, using E DENV-2 and E DENV-3 protein as their backbones, which could give immune response toward four dengue virus serotype (tetravalent). The specific objective of this research is to design and predict the tertiary structure of dengue virus tetravalent vaccines by homology modeling method and determine which vaccines that have the best functionality according to its tertiary structure^[11-14].

MATERIALS AND METHODS

Searching for E DENV-2 and E DENV-3 protein PDB data as the backbone: E DENV-2 and E DENV-3 protein data can be downloaded from PDB database in the RSCB Protein Data Bank, from the <http://www.rcsb.org/pdb/> by entering PDB-ID code of 1OAN for E DENV-2 protein and 1UZG for E DENV-3 protein. The PDB files were used as backbones.

Searching for serotype sequences of DENV1-4: The search for E protein sequences of four dengue virus serotypes (DENV1-4) was done by accessing Viral Bioinformatics Resource Center Canada (VBRCCa) in their website at <http://www.athena.bioc.uvic.ca/>, using computer with internet access. The sequences format that utilized for saving downloaded E protein sequences in each DENV serotypes was FASTA format. This format was submitted for query of further analysis.

T cell epitope of E DENV protein prediction: T cell epitope prediction was done with MULTIPRED server by using Hidden Markov Model (HMM) and Artificial Neural Network (ANN) algorithm. The server website is <http://antigen.i2r.a-star.edu.sg/multipred/>. MULTIPRED needs data input of E DENV 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 and it was conducted for each DENV1-4 serotype. The chosen interface is sorted result.

B cell epitope of E DENV-2 and E DENV-3 protein prediction: B cell epitope prediction was done by using Conformational Epitope Prediction (CEP) server, which can be accessed freely in their website at <http://202.41.70.74:8080/cgi-bin/cep.pl>. CEP server needs protein PDB data as its input for B cell epitope prediction. The prediction result was peptide sequences with their following position in the E DENV protein sequences and their accessibility.

Dengue vaccine peptide sequences determination: Epitope substitution of E DENV protein sequences was done using Notepad text editor. The first substitution used E DENV-2 protein and the second was E DENV-3 protein as their backbone. The substitution of three epitope position will be resulted in one new E DENV protein sequence. First vaccine type would be utilized by DENV serotype 1, 3 and 4 as its backbone. Moreover, Second vaccine type would be utilized by DENV serotype 1, 2 and 4 as its backbone. This whole procedure resulted in new Tetravalent Dengue vaccine peptides design in FASTA format.

Dengue vaccine peptide sequences comparison with genbank database: Comparison of dengue vaccine peptide sequences with protein database in the GenBank was proceeded by BLAST toolbox on the National Center for Biotechnology Information website at <http://www.ncbi.nlm.nih.gov>. Data input for these comparisons were dengue vaccine peptide sequences, while the database was in PDB format.

Template inquiry: The E DENV-2 and E DENV-3 protein were visualized by using Deep View/Swiss Pdb-Viewer 3.7 (SP5) and Swiss Model server, which could be freely accessed at <http://www.expasy.org/swissmod/SWISS-MODEL.html>. The residue in the sequences could be seen at control panel menu on the Deep View. Moreover, Inquiry of PDB template for homology modeling was executed, based on sequences similarity with Peptide vaccine. The inquiry was done by Swiss Model server.

Homology modeling: Dengue vaccine peptide structure prediction was accomplished by homology modeling method, using Deep View/Swiss Pdb-Viewer 3.7 (SP5) and Swiss Model server.. Structure prediction of dengue vaccine peptides was proceeded by Optimise mode

Structure of dengue vaccine peptides prediction with optimise mod: Dengue vaccine peptides were visualized as straight chain by using Deep View/Swiss-Pdb Viewer software, by accessing Load Raw Sequence to Model menu. Additionally, the PDB template was visualized. After the visualization, the vaccine sequences and the templates were bind by selecting Magic Fit menu in the Deep View software. The aligned residues could be seen on the Alignment menu at Deep View. The visualization of dengue vaccine peptides tertiary structure with the binding result was accomplished by eliminating the tertiary structure template, so the software only shown the tertiary structure of the dengue vaccine peptides.

Reparation of dengue vaccine peptides structure: The reparation of dengue vaccine peptides structure was done by Deep View Software. First step was repairing the overlapping residue by using the Fix Selected Side chains: Quick and Dirty menu. Second step was utilizing Swiss Model server. The repaired Dengue Vaccine Peptides with Swiss-Pdb Viewer were sent to Swiss Model server in the PDB format. Moreover, the structure optimisation was conducted by choosing the Optimise Mode menu in the Swiss Model server.

Visualization and evaluation of dengue vaccine peptides tertiary structure homology modeling result: The homology modeling result of the sequences were visualized by Swiss-Pdb viewer. In order to enhance the appearance of the visualization, interface manipulation was carried. The evaluation of dengue vaccine peptides tertiary structure was done by finding existing overlapping residues and by analyzing the Ramachandran Plot and dengue vaccine peptide structure with Ramachandran Plot menu in the Swiss-Pdb viewer.

Comparison of dengue vaccine peptide tertiary structure with genbank database: The comparison of dengue vaccine peptide structure with protein data in the PDB format on GenBank, was done using VAST toolbox of National Center for Biotechnology Information (NCBI) at <http://www.ncbi.nlm.nih.gov/Structure/VAST/vastsearch.html> with entering the PDB structural data of prediction result.

RESULTS

DENV intra serotypes that utilized for this dengue vaccine peptides design were acquired from previous research of our group. One representative was chosen from each DENV serotype, based on the multiple alignment result, incidence level and geographical area. The chosen DENV intra serotypes are:

DENV-1 → DENV-1-SIN-8114_93
DENV-2 → DENV-2-ThNH28_93
DENV-3 → DENV-3-SIN-8120_95
DENV-4 → DENV-4-SIN8976_95

The prediction result of T cell epitope position with HMM and ANN algorithm method, Binding Score and DR HLA class II with the highest rank based on MULTIPRED prediction will be shown in Table 1-3.

The substituted T cell epitope of E DENV-2 and DENV-3 protein was determined by comparing the non binder amino acids of MULTIPRED prediction with the conformational epitope amino acids of CEP prediction. If the non binder amino acids of E DENV-2 and DENV-3 MULTIPRED prediction are the same amino acids for constituting the B cell epitope by CEP prediction, then the non binder amino acids would not selected for substitution with high binder amino acids from MULTIPRED prediction. The B cell epitope could be able to induce specific antibody^[16-18]. The non binding peptides of E DENV-2 and DENV-3 protein will be shown in Table 4 and 5.

Table 1: Epitope Binding of E DENV protein HMM method

Serotype	DENV1	DENV2	DENV4
Peptide	VFGTAYGVL	WLVRHQWFL	WLVHKQWFL
Position	439-447	206-214	206-214
Score	8.27	8.40	8.23

Table 2: Epitope Binding of E DENV protein ANN method

Serotype	DENV1	DENV2
Peptide	FCVRRTFVD	IQMSSGNLL
Position	90-98	270-278
Score	6.36	7.04

Table 3: Epitope Binding of E DENV protein ANN method

Serotype	DENV3	DENV4
Peptide	YVCKHTYVD	YRGAKRMAI
Position	90-98	406-414
Score	6.73	6.59

Table 4: Non binding peptides of E DENV-2 protein

	TKDFE	QPENL	PVNIE
Peptide	LIKT	EYTI	AEPP
Position	40-48	131-139	364-372

Table 5: Non binding peptides of E DENV-3 protein

	RQWFFD	NKPTLD	GTILIKV
Peptide	LPL	IEL	EY
Position	208-216	37-45	316-324

```
>ANN1_vaccine
MRCIGISNRDFVEGVSGGSWVDIVLEHGSCVTT
MAKNKPFVCRRTFVDEAKQPATLRKYCIEAKL
TNTTTESRCPQTQGEPTLNEEQDKRFVCKHSMV
DRGWGNGCGLFGKGGIVTCAMFTCKKNMEGK
IVYVCKHTYVDVTPHSGEEHAVGNDTGKHHG
KEVKITPQSSSITEAELTGYGTVTMECSPTGLDF
NEMVLLQMKDKAWLVHRQWFLDPLPWLPG
ADTQGSNWIQKETLVTFKNPHAKKQDVVVLG
SQEGAMHTALTGATEIQMSSGNLLFTGHLKCR
LRMDKQLKGMYSYMCCTGKFKVVKEIAETQH
GTIVIRVQYEGDGSCKIPFEIMDLEKRHVLRGL
ITVNPVIVTEKDSYRGAKRMAIFGDSYIIIGVEPG
QLKLNWFKK
```

Fig. 1: ANN1 vaccine

```
>ANN2_vaccine
MRCIGISNRDFVEGVSGGSWVDIVLEHGSCVTT
MAKNKPFVCRRTFVDEAKQPATLRKYCIEAKL
TNTTTESRCPQTQGEPTLNEEQDKRFVCKHSMV
DRGWGNGCGLFGKGGIVTCAMFTCKKNMEGK
IVYRGAKRMAIVVTPHSGEEHAVGNDTGKHHG
KEVKITPQSSSITEAELTGYGTVTMECSPTGLDF
NEMVLLQMKDKAWLVHRQWFLDPLPWLPG
ADTQGSNWIQKETLVTFKNPHAKKQDVVVLG
SQEGAMHTALTGATEIQMSSGNLLFTGHLKCR
LRMDKQLKGMYSYMCCTGKFKVVKEIAETQH
GTIVIRVQYEGDGSCKIPFEIMDLEKRHVLRGL
ITVNPVIVTEKDSYVCKHTYVDFGDSYIIIGVEPG
QLKLNWFKK
```

Fig. 2: ANN2 vaccine

The substitution was conducted with three epitopes. The different substituted epitope position was resulted in six dengue vaccine peptides for ANN vaccine and the same amount for the HMM vaccine.

```
>ANN3_vaccine
MRCIGISNRDFVEGVSGGSWVDIVLEHGSCVTT
MAKNKPYVCKHTYVDEAKQPATLRKYCIEAK
LTNTTTESRCPQTQGEPTLNEEQDKRFVCKHSM
VDRGWGNGCGLFGKGGIVTCAMFTCKKNMEG
KIVFVCRRTFVDVTPHSGEEHAVGNDTGKHHG
KEVKITPQSSSITEAELTGYGTVTMECSPTGLDF
NEMVLLQMKDKAWLVHRQWFLDPLPWLPG
ADTQGSNWIQKETLVTFKNPHAKKQDVVVLG
SQEGAMHTALTGATEIQMSSGNLLFTGHLKCR
LRMDKQLKGMYSYMCCTGKFKVVKEIAETQH
GTIVIRVQYEGDGSCKIPFEIMDLEKRHVLRGL
ITVNPVIVTEKDSYRGAKRMAIFGDSYIIIGVEPG
QLKLNWFKK
```

Fig. 3: ANN3 vaccine

```
>ANN4_vaccine
MRCIGISNRDFVEGVSGGSWVDIVLEHGSCVTT
MAKNKPYVCKHTYVDEAKQPATLRKYCIEAK
LTNTTTESRCPQTQGEPTLNEEQDKRFVCKHSM
VDRGWGNGCGLFGKGGIVTCAMFTCKKNMEG
KIVYRGAKRMAIVVTPHSGEEHAVGNDTGKHHG
GKEVKITPQSSSITEAELTGYGTVTMECSPTGL
DFNEMVLLQMKDKAWLVHRQWFLDPLPWLPG
ADTQGSNWIQKETLVTFKNPHAKKQDVVVLG
LGSQEGAMHTALTGATEIQMSSGNLLFTGHLK
CRLRMDKQLKGMYSYMCCTGKFKVVKEIAET
QHGTIVIRVQYEGDGSCKIPFEIMDLEKRHVLR
GRLITVNPVIVTEKDSFVCRRTFVDFGDSYIIIGVE
PGQLKLNWFKK
```

Fig. 4: ANN4 vaccine

```
>ANN5_vaccine
MRCIGISNRDFVEGVSGGSWVDIVLEHGSCVTT
MAKNKPYRGAKRMAIEAKQPATLRKYCIEAKL
TNTTTESRCPQTQGEPTLNEEQDKRFVCKHSMV
DRGWGNGCGLFGKGGIVTCAMFTCKKNMEGK
IVFVCRRTFVDVTPHSGEEHAVGNDTGKHHG
EVKITPQSSSITEAELTGYGTVTMECSPTGLDFN
EMVLLQMKDKAWLVHRQWFLDPLPWLPGA
DTQGSNWIQKETLVTFKNPHAKKQDVVVLGS
QEGAMHTALTGATEIQMSSGNLLFTGHLKCR
LRMDKQLKGMYSYMCCTGKFKVVKEIAETQH
GTIVIRVQYEGDGSCKIPFEIMDLEKRHVLRGLI
TVNPVIVTEKDSYVCKHTYVDFGDSYIIIGVEPG
LKLNFVFKK
```

Fig. 5: ANN5 vaccine

```
>ANN6_vaccine
MRCIGISNRDFVEGVSGGSWVDIVLEHGSCVTT
MAKNKPYRGAKRMAIEAKQPATLRKYCIEAKL
TNTTTESRCPQTQGEPTLNEEQDKRFVCKHSMV
DRGWGNGCGLFGKGGIVTCAMFTCKKNMEGK
IVVCKHTYVDVTPHSGEEHAVGNDTGKHHG
EVKITPQSSSITEAELTGYGTVTMECSPTGLDFN
EMVLLQMKDKAWLVHRQWFLDPLPWLPGA
DTQGSNWIQKETLVTFKNPHAKKQDVVVLGS
QEGAMHTALTGATEIQMSSGNLLFTGHLKCR
LRMDKQLKGMYSYMCCTGKFKVVKEIAETQH
GTIVIRVQYEGDGSCKIPFEIMDLEKRHVLRGLI
TVNPVIVTEKDSFVCRRTFVDFGDSYIIIGVEPG
LKLNFVFKK
```

Fig. 6: ANN6 vaccine

The total twelve dengue vaccine peptides design was constructed by MULTIPRED server. The vaccines were shown in Fig. 1-12.

```
> HMM1_Vaccine
MRCVGVGNRDFVEGLSGATWVDVVLEHGGC
VTTMAKVFGTAYGVLQKTEATQLATLRKLCIE
GKITNITDTSRCPQTQGEAILPEEQDQNYVCKHT
YVDRGWGNGCGLFGKGSVLTCAKFCLESIEG
KIVQHENLKYTVIITVHTGDQHVGNETQGV
TAEITSQASTAEAILPEYGTGLGLECSPRTGLDFNE
MILLTMKDKAWMVHWLVHRQWFLPWTSGAT
TKTPTWNRKELLVTFKNAHAKKQEVVVLGSQ
EGAMHTALTGATEIQTSGGTSIFAGHLKCRLK
MDKLLKLGMSYAMCLNTFVLKKEVSETQHWL
VHKQWFLKGEDAPCKIPFSTEDGQKKAHNGRL
ITANPVVTKKEEPVNIEAEPFFGESNIVIGIDKA
LKINWYRK
```

Fig. 7: HMM1 vaccine

```
> HMM5_Vaccine
MRCVGVGNRDFVEGLSGATWVDVVLEHGGC
VTTMAKWLVHKQWFLQKTEATQLATLRKLCI
EGKITNITDTSRCPQTQGEAILPEEQDQNYVCKH
TYVDRGWGNGCGLFGKGSVLTCAKFCLESIE
GKIVQHENLKYTVIITVHTGDQHVGNETQGV
TAEITSQASTAEAILPEYGTGLGLECSPRTGLDFN
EMILLTMKDKAWMVHVFGTAYGVLPWTSGAT
TKTPTWNRKELLVTFKNAHAKKQEVVVLGSQ
EGAMHTALTGATEIQTSGGTSIFAGHLKCRLK
MDKLLKLGMSYAMCLNTFVLKKEVSETQHWL
VHRQWFLKGEDAPCKIPFSTEDGQKKAHNGRL
ITANPVVTKKEEPVNIEAEPFFGESNIVIGIDKA
LKINWYRK
```

Fig. 11: HMM5 vaccine

```
> HMM2_Vaccine
MRCVGVGNRDFVEGLSGATWVDVVLEHGGC
VTTMAKVFGTAYGVLQKTEATQLATLRKLCIE
GKITNITDTSRCPQTQGEAILPEEQDQNYVCKHT
YVDRGWGNGCGLFGKGSVLTCAKFCLESIEG
KIVQHENLKYTVIITVHTGDQHVGNETQGV
TAEITSQASTAEAILPEYGTGLGLECSPRTGLDFNE
MILLTMKDKAWMVHWLVHKQWFLPWTSGAT
TKTPTWNRKELLVTFKNAHAKKQEVVVLGSQ
EGAMHTALTGATEIQTSGGTSIFAGHLKCRLK
MDKLLKLGMSYAMCLNTFVLKKEVSETQHWL
VHRQWFLKGEDAPCKIPFSTEDGQKKAHNGRL
ITANPVVTKKEEPVNIEAEPFFGESNIVIGIDKA
LKINWYRK
```

Fig. 8: HMM2 vaccine

```
> HMM6_Vaccine
MRCVGVGNRDFVEGLSGATWVDVVLEHGGC
VTTMAKWLVHKQWFLQKTEATQLATLRKLCI
EGKITNITDTSRCPQTQGEAILPEEQDQNYVCKH
TYVDRGWGNGCGLFGKGSVLTCAKFCLESIE
GKIVQHENLKYTVIITVHTGDQHVGNETQGV
TAEITSQASTAEAILPEYGTGLGLECSPRTGLDFN
EMILLTMKDKAWMVHWLVHRQWFLPWTSGA
TTKTPTWNRKELLVTFKNAHAKKQEVVVLGS
QEGAMHTALTGATEIQTSGGTSIFAGHLKCRLK
MDKLLKLGMSYAMCLNTFVLKKEVSETQHVF
GTAYGVLKGEDAPCKIPFSTEDGQKKAHNGRL
ITANPVVTKKEEPVNIEAEPFFGESNIVIGIDKA
LKINWYRK
```

Fig. 12: HMM6 vaccine

```
> HMM3_Vaccine
MRCVGVGNRDFVEGLSGATWVDVVLEHGGC
VTTMAKWLVHRQWFLQKTEATQLATLRKLCI
EGKITNITDTSRCPQTQGEAILPEEQDQNYVCKH
TYVDRGWGNGCGLFGKGSVLTCAKFCLESIE
GKIVQHENLKYTVIITVHTGDQHVGNETQGV
TAEITSQASTAEAILPEYGTGLGLECSPRTGLDFN
EMILLTMKDKAWMVHVFGTAYGVLPWTSGAT
TKTPTWNRKELLVTFKNAHAKKQEVVVLGSQ
EGAMHTALTGATEIQTSGGTSIFAGHLKCRLK
MDKLLKLGMSYAMCLNTFVLKKEVSETQHWL
VHKQWFLKGEDAPCKIPFSTEDGQKKAHNGRL
ITANPVVTKKEEPVNIEAEPFFGESNIVIGIDKA
LKINWYRK
```

Fig. 9: HMM3 vaccine

```
> HMM4_Vaccine
MRCVGVGNRDFVEGLSGATWVDVVLEHGGC
VTTMAKWLVHRQWFLQKTEATQLATLRKLCI
EGKITNITDTSRCPQTQGEAILPEEQDQNYVCKH
TYVDRGWGNGCGLFGKGSVLTCAKFCLESIE
GKIVQHENLKYTVIITVHTGDQHVGNETQGV
TAEITSQASTAEAILPEYGTGLGLECSPRTGLDFN
EMILLTMKDKAWMVHWLVHKQWFLPWTSGA
TTKTPTWNRKELLVTFKNAHAKKQEVVVLGS
QEGAMHTALTGATEIQTSGGTSIFAGHLKCRLK
MDKLLKLGMSYAMCLNTFVLKKEVSETQHVF
GTAYGVLKGEDAPCKIPFSTEDGQKKAHNGRL
ITANPVVTKKEEPVNIEAEPFFGESNIVIGIDKA
LKINWYRK
```

Fig. 10: HMM4 vaccine

Table 6: Overlapping Residues comparison before treatment and after the treatment in E DENV2 ANN1-3 vaccines

Vaccine type	ANN1	ANN2	ANN3
Before quick and dirty treatment	21	21	25
After quick and dirty treatment	5	3	9

Table 7: Overlapping Residues comparison before treatment and after the treatment in E DENV2 ANN4-6 vaccines

Vaccine type	ANN4	ANN5	ANN6
Before quick and dirty treatment	19	24	22
After quick and dirty treatment	6	5	6

Table 8: Overlapping Residues comparison before treatment and after the treatment in E DENV3 HMM1-3 vaccines

Vaccine type	HMM1	HMM2	HMM3
Before quick and dirty treatment	23	23	19
After quick and dirty treatment	8	8	6

The overlapping residues in the protein were assessed in the layers Infos menu in Deep View. They were repaired by Fix Selected Side chains: Quick and Dirty menu in the software. The overlapping residues were decreased significantly after the treatment. The overlapping residues in both types of vaccines were still existed, as shown in the Table 6-9.

The PDB data from Optimise Mode SwissModel server were visualized with Deep View software as shown in Fig. 13-24.

Table 9: Overlapping Residues comparison before treatment and after the treatment in E DENV3 HMM4-6 vaccines

Vaccine type	HMM4	HMM5	HMM6
Before quick and dirty treatment	23	16	20
After quick and dirty treatment	8	6	8

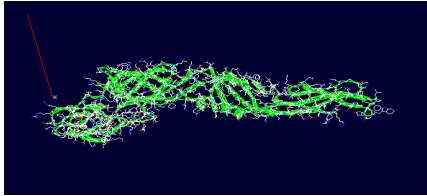


Fig. 13: ANN1 vaccine

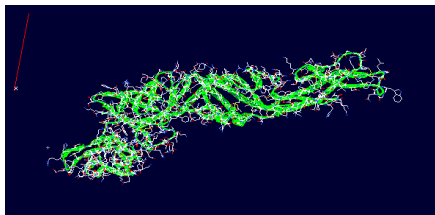


Fig. 14: ANN2 vaccine

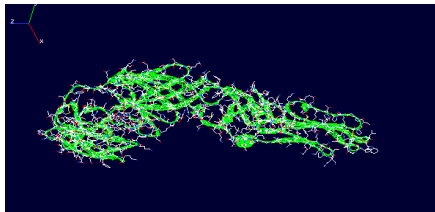


Fig. 15: ANN3 vaccine

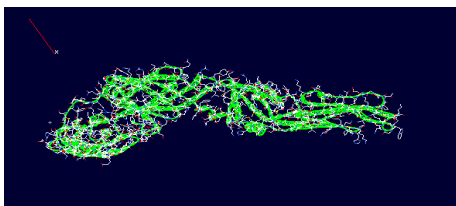


Fig. 16: ANN4 vaccine

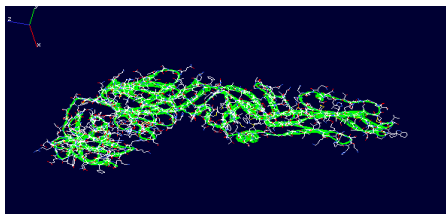


Fig. 17: ANN5 vaccine

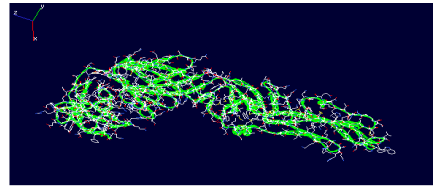


Fig. 18: ANN6 vaccine

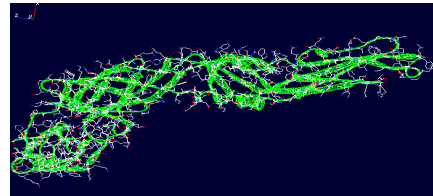


Fig. 19: HMM1 vaccine

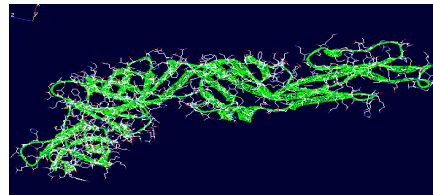


Fig. 20: HMM2 vaccine

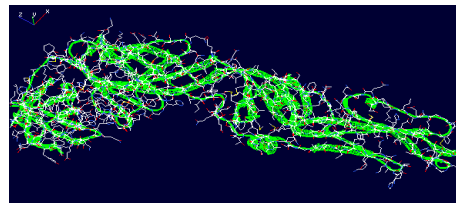


Fig. 21: HMM3 vaccine

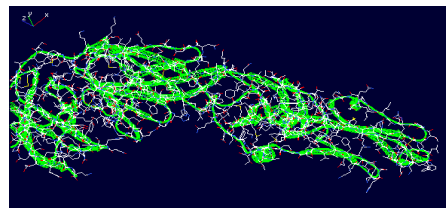


Fig. 22: HMM4 vaccine

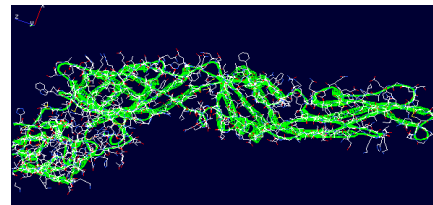


Fig. 23: HMM5 vaccine

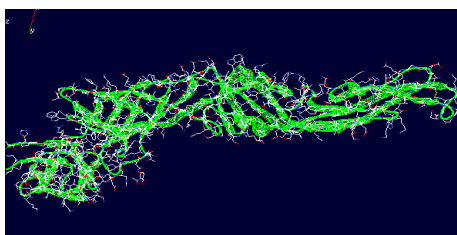


Fig. 24: HMM6 vaccine

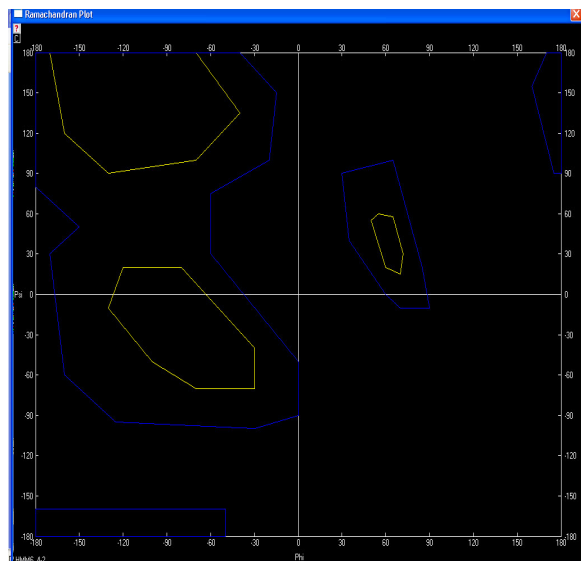


Fig. 25: Sample of Ramachandran Plot

The peptide structure evaluation was accomplished by assessing the total amount of overlapping residues and computing the Ramachandran Plot of the vaccine. The result of the evaluation was no more overlapping residues in the vaccines. The example of Ramachandran plot was shown in Fig. 25.

The glycine plot may be allowed in the disallowed region. Glycine could make the unlimited amount of Φ and Ψ angle. The non glycine residues in the disallowed region were showing the stability of protein structure. If the percentage was more than 15% of total protein residue, then the protein was unstable. The non glycine residues plot in the disallowed region will be shown in Table 10-13.

The tertiary structure of dengue peptide vaccine was compared with the database of Vector Alignment Search Tool (VAST), which was integrated in the National Center for Biotechnology Information website. The database could be accessed in the following website <http://www.ncbi.nlm.nih.gov/Structure/VAST/vastsearch.html>. The VAST result of ANN and HMM vaccines were shown in Table 14-17.

Table 12: The non glycine residues plot in the disallowed region of the HMM1-3 vaccine protein structure

Vaccine Type	HMM1	HMM2	HMM3
Total residues in the disallowed region	8	8	8

Table 13: The non glycine residues plot in the disallowed region of the HMM4-6 vaccine protein structure

Vaccine Type	HMM4	HMM5	HMM6
Total residues in the disallowed region	8	8	8

Table 14: VAST result of ANN1-3 vaccines

Parameter	ANN1 vaccine	ANN2 vaccine	ANN3 vaccine
score	64.7	62.8	64.7
RMSD (Å)	0.1	0.1	0.1
Percent identity	91.1	91.1	90.9

Table 15: VAST result of ANN4-6 vaccines

Parameter	ANN4 vaccine	ANN5 vaccine	ANN6 vaccine
score	64.7	64.7	64.7
RMSD (Å)	0.1	0.1	0.1
Percent identity	91.1	91.1	91.4

Table 16: VAST result of HMM1-3 vaccines

Parameter	HMM1 vaccine	HMM2 vaccine	HMM3 vaccine
Score	62.7	62.7	62.7
RMSD (Å)	0.1	0.1	0.1
Percent identity	93.9	93.9	93.6

Table 17: VAST result of HMM4-6 vaccines

Parameter	HMM4 vaccine	HMM5 vaccine	HMM6 vaccine
Score	64.5	62.7	64.5
RMSD (Å)	0.1	0.1	0.1
Percent identity	93.6	93.6	93.6

DISCUSSION

This dengue virus tetravalent vaccine design was done by in silico method. The method was using E DENV-2 protein and DENV-3 as their backbones. The reason for choosing E DENV protein was because of its function on viral attachment at the host cell surface and for facilitating immune response at the host cell. E DENV-2 and E DENV-3 proteins were chosen as the backbones because of their high prevalence in the south east Asian region^[8,19].

Vaccine design was accomplished by using homology modeling method. This method needs three dimensional protein molecule structure in the PDB format. PDB files could be downloaded from Protein Data Bank, which supervised by Research Collaboratory for Structural Biology in their websites <http://www.rcsb.org/pdb> and could be accessed freely by internet access.

The E DENV-2 and DENV-3 files had been found. They are 1OAN.pdb and 1UZG.pdb. The files could be

visualized with molecular viewer software Deep View and then both files were applied in Homology modeling method as the templates for the vaccine backbone. The downloading of the PDB files will provide E DENV-2 and E DENV-3 protein in FASTA format, with 392 amino acids and 394 amino acids respectively.

The E DENV1-4 protein sequences were downloaded from Virus Bioinformatics Resource Center Canada (VBRCCa) by internet access to the website <http://athena.bioc.uvic.ca>. This website provided genome information of all kind of virus. Dengue virus data sequences were classified as RNA single strand, Flaviviridae family and Flavivirus genus. Sequences database of dengue virus E protein was saved in FASTA format with different definition line. The definition line of this website was bolded without the ">" mark, which useful for underlining the DENV strain name.

With the SIN abbreviation is for Singapore and Th is Thailand. The chosen intra serotype is prevalent and causing epidemic in the south East Asian region^[19].

The T cell epitope of E DENV-2 and E DENV-3 protein result were applied for peptide substitution determination. The non binding peptides of E DENV-2 from ANN method were substituted with high binder peptides of E DENV protein from E DENV1, 3 and 4 proteins. Furthermore, the non binding peptides of E DENV-3 from HMM method were substituted with high binder peptides of E DENV protein form E DENV 1, 2 and 4 proteins. The substitution of the non binder peptides must not overlap with the B cell epitope.

T cell epitope prediction of E DENV1-4 protein was conducted with MULTIPRED server at their website <http://antigen.i2r.a-star.edu.sg/multipred/>. This toolbox predicted the potential antigenic determinant peptides or epitopes for Human Leukocyte Antigen (HLA) or Major Histocompatibility Complex (MHC) binding. The binding is prerequisite for T cell receptor recognition. The necessary data input is amino acid sequences of E DENV protein in FASTA format, without the sequence identity line^[16].

MULTIPRED has two algorithms methods that can be utilized for T cell epitope prediction; they are Hidden Markov Model (HMM) and Artificial Neural Network (ANN). This research applied both algorithms for epitopes prediction. The E DENV-2 backbone substitution was using ANN method and E DENV-3 backbone substitution was using HMM method. Those algorithms have insignificant difference in terms of analysis result^[15,20].

The chosen parameter of MULTIPRED server was the election of HLA molecule class for epitope prediction. HLA was consisted of two classes, class I

for intracellular peptide fragment binding which activated by CD8⁺; and class II for extra cellular peptide fragment binding which activated by CD4⁺. DR HLA class II was chosen because the dengue vaccine peptides are extra cellular peptide vaccines (exogenous antigen, not endogenous antigen like attenuated virus). The other reason is the increasing amount of CD4⁺ serum in the human patients during the dengue infection^[15,20].

The output T cell epitope of E DENV protein is peptide sequences with different binding score. The evaluation of binding score shows the strength of epitope binding with HLA. If the binding between epitope and HLA is strong, then the possibility of T cell recognition toward the epitope is high. The amino acid sequences with high binding score would have a high possibility of inducing antibody response. The given MULTIPRED score has a range of 1-9. Peptides with score of 4-9 were predicted as binders (8-9 is high binder, 6-7 is moderate binder and 4-5 is low binder). Moreover, the score 1-3 were predicted as non binders or not epitopes. The analysis was accomplished by scanning every 30 amino acids of E DENV protein, starting from the first amino acids until the last ones.

B cell epitope of E DENV protein was predicted with Conformational Epitope Prediction (CEP) server, which could be accessed with internet connection at the website <http://202.41.70.74:8080/cgi-bin/cep.pl>. It was managed by bioinformatics institute in Pune University, India. CEP server needs data input in PDB-ID code or downloaded PDB file. The B cell epitope prediction was done with E DENV-2 and E DENV-3 proteins, because only both proteins have PDB-ID entry. Both proteins were utilized as the vaccine backbones^[15].

The B cell epitope prediction of E DENV-2 protein with CEP server was showing the positions of epitopes, but no binding score with the antibody. During the dengue vaccine peptides determination, the existence of the B cell epitope prediction was conserved.

The similarity of dengue vaccine peptides design with native E DENV-2 and DENV-3 protein could be verified by comparing the vaccines sequences with the Protein Data Bank database. This comparison could be done by using BLASTp toolbox in the website <http://www.ncbi.nlm.nih.gov>.

The BLAST comparison result of the dengue vaccine peptides is 91% with E DENV-2 protein and 93% with E DENV-3 protein, for all twelve vaccine sequences. This result has shown that the vaccine sequences have high degree of similarity with native E DENV-2 and DENV-3 protein. The sequence was considered similar if the percent identity higher than

25%. The high percent identity for the vaccine sequences was caused by the small number of substituted native E DENV-2 and DENV-3 protein, so the vaccine sequences were very similar with the native E protein sequences.

The tertiary structure prediction of dengue vaccine peptides with homology modeling could be done if the sequence similarity between the vaccine sequences and the native protein sequences in the database was at least 20% identity. The BLASTp comparison result between vaccine sequences and protein sequences in the database have shown a high result of 91% for E DENV-2 protein vaccine backbone and 93% for E DENV-3 protein vaccine backbone. This result made the homology modeling method for tertiary structure prediction possible^[20].

Homology modeling was carried out with Swiss Model server in their website <http://www.expasy.org/swissmodel/SWISS-MODEL.html/>. This research was applying Optimise mode for predicting the 3-D structure of the proteins.

The dengue vaccine peptide sequences in the FASTA format were visualized in the straight chain form, by choosing the Load Raw Sequence to Model menu in the Deep View software. The objective of this visualization was to search for the vaccine peptide template, because the dengue vaccine peptide visualization made direct template searching to the SWISS MODEL server possible. The templates were three dimensional protein structures in the PDB format, which have sequences similarity with the vaccine sequences.

The template searching of the twelve vaccine sequences in the ExPDB library was resulted in several templates. The template for E DENV-2 protein was 1OANA.pdb, while the template for E DENV-3 protein was 1UZGA.pdb and 1UZGB.pdb. The result was come to our expectation, because the templates were PDB data for E DENV-2 and DENV-3 native protein.

Tertiary structure prediction was done by choosing the optimise mode. Vaccine sequences were visualized in the straight chain and the templates were visualized with the Deep View software. The sequence constituent residues of the vaccines and the templates could be observed by accessing Alignment and Control Panel menu. Moreover, the vaccine sequences and the templates were attached with the Magic Fit menu in the software. The attachment was based on the amino acids similarity, in order to create the same folding. This attachment has turned the straight chain vaccine sequences into the tertiary structure.

The tertiary structure of dengue vaccine peptide sequences could be visualized with Deep View, by

making the template structure invisible. This action could be accomplished by accessing Control Panel or Layers Infos menu in the software.

The Magic Fit attachment of Deep View was done based on the assumption, that the same amino acids sequences will produce the same protein folding. In the reality, there are differences of protein folding in the same amino acids, because the significant role of amino acids residue in it.

The tertiary structure of the vaccines was submitted to SwissModel server with Optimise mode. The peptide structures were repaired totally in the server. The reparation result was sent with email in the PDB files^[21-23].

The Ramachandran Plot of the twelve dengue vaccine peptides was evaluated by accessing the Ramachandran Plot menu in Deep View software. Every protein residue has one Φ angle as the x-axis and Ψ angle as the y-axis. This arrangement of every single residue was represented as one plot. The Φ angle was a dihedral angle of N-C α bond, while the Ψ angle was a dihedral angle of C α -C bond^[24]. This plot has an area bordered with blue line, which was the coordinate of protein secondary structure area. This area was the allowed region. Moreover, the disallowed region was beyond the allowed region. The non glycine residue was forbidden in the disallowed region.

The plot was arranged for decreasing the total residues in the disallowed region. The Ramachandran Plot of both types of the vaccines has shown that the non glycine residues were few. This result has shown that the total amount of residues in the disallowed region were fewer than 15% of total protein residue. It could be inferred from the plot that the dengue vaccine peptides were having excellent stability. The result showed that the E native DENV2 and E DENV3 have high degree of similarity with ANN and HMM vaccines, respectively. VAST toolbox was comparing the tertiary structure of dengue vaccine with native E DENV2 and E DENV3, by superimposing between the structures^[25].

The results of percent identity entry in the whole vaccines were higher than 90%. The percent identity entries of the vaccines were identical with the BLAST results, which was also higher than 90%. The results of RMSD entry for all vaccines were 0.1 Å. The VAST result 0.1 Å means that the vaccines structures were essentially indistinguishable with the native E DENV protein structure.

The VAST score of whole vaccines were diverged. The ANN1,3-6 vaccines have the highest VAST score among the ANN vaccines, while the HMM 4 and 6 have the highest score among the HMM vaccines. It

could be inferred from the VAST scores, that ANN1, 3-6 are the best vaccines among the ANN; while the HMM4 and 6 are the best vaccines among the HMM.

CONCLUSION

This research has invented twelve dengue vaccines peptides. They are ANN1, ANN2, ANN3, ANN4, ANN5, ANN6, HMM1, HMM2, HMM3, HMM4, HMM5 and HMM6. The BLAST and VAST score results of the vaccines have shown more than 90% identity with native E DENV2 and DENV3 protein. The Ramachandran plot of the vaccines have shown that less than 15% of protein residues were inside the disallowed region. The determining factor for the quality of the vaccines came from the VAST score, which showed that ANN1, 3-6 were the best vaccines among the ANN vaccines, while HMM4 and 6 were the best vaccines among the HMM vaccines. This research could be applied for the wet laboratory and computerised vaccine design.

ACKNOWLEDGEMENT

The researchers are grateful to Ridla Bakri PhD, Chairman of Department of Chemistry, Faculty of Science, University of Indonesia for his support of this research. This Research was supported by grants from Ministry of Education (Grant Hibah Bersaing).

REFERENCES

1. Chatuverdi, U.C., R. Shrivastava and R. Nagar, 2005. Dengue vaccines: Problems and prospects. *Rev. Art. Indian J. Med. Res.*, 121: 639-652. <http://www.ncbi.nlm.nih.gov/pubmed/15937367>.
2. WHO, 2007. SEARO WHO report. <http://www.searo.who.int>.
3. Guglani, L. and S.K. Kabra, 2005. T cell immunopathogenesis of dengue virus infection. *Dengue Bull.*, 29:58-69. <http://direct.bl.uk/bld/PlaceOrder.do?UIN=186391443&ETOC=RN&from=searchengine>.
4. Kusriastuti, R. and S. Sutomo, 2005. Evolution of dengue prevention and control programme in Indonesia. *Dengue Bull.*, 29: 1-7. [http://www.searo.who.int/LinkFiles/Dengue_Bulletins_Volumes_29_\(2005\)_CHAPTER01.pdf](http://www.searo.who.int/LinkFiles/Dengue_Bulletins_Volumes_29_(2005)_CHAPTER01.pdf).
5. IVI Team, 2007. International Vaccine Institute. <http://www.ivi.org>.
6. Elgert, K.D., 1996. *Immunology: Understanding the Immune System*. 2nd Edn., Wille-Liss, Inc., New York, ISBN: 978-0-471-11680-6, pp: 480.
7. Halstead, S.B. and J. Deen, 2002. The future of dengue vaccine. *Lancet*, 360: 1243-1245. <http://linkinghub.elsevier.com/retrieve/pii/S0140673602112761>.
8. Gancalves, A.P., R. Purcell, C.J. Lai, 2004. Epitope determinants of chimpanzee FAB antibody that efficiently cross neutralizes dengue type 1 and 2 viruses map to inside and in close proximity to fusion loop of the dengue type2 virus envelope glycoprotein. *J. Virol.*, 78: 12919-12928. DOI: 10.1128/JVI.78.23.12919-12928.2004.
9. Kols, A. and J. Sehrris, 2000. HPV vaccines: Promise and challenges. <http://www.ponline.org/docs/1710/307320.html>.
10. Thomson, S.A., S.R. Burrows, I.S. Misko, D.J. Moss, B.E.H. Coupar and R. Khanna, 1998. Targeting for polypeptide protein incorporating multiple class II-restricted viral epitopes to the secretory pathway facilitates immune recognition by CD4+ cytotoxic T lymphocytes: A novel approach to vaccine design. *J. Virol.*, 72: 2246-2252. <http://jvi.asm.org/cgi/content/abstract/72/3/2246>.
11. Baxevanis, A.D. and B.F.F. Quellerie, 2004. *Bioinformatics, a Practical Guide to the Analysis of Genes and Proteins*. 3rd Edn., John Wiley and Sons, Inc., USA., ISBN: 10: 0471478784, pp: 560.
12. Claverie, J.M. and C. Notredame, 2003. *Bioinformatics for Dummies*. 1st Edn., Willey Publishing Inc., New York, ISBN: 0764516965, pp: 480.
13. Modis, Y., S. Ogata, D. Clements and S. Harrison, 2003. A ligand-binding pocket in the dengue virus envelope glycoprotein. *Proc. Natl. Acad. Sci. USA.*, 100: 6986-6991. DOI: 10.1073/pnas.0832193100.
14. VBRC Team, 2008. Viral bioinformatics resource center. <http://athena.bioc.uvic.ca>.
15. Kulkarni-Kale, U., S. Bhosle and A.S. Kolaskar, 2005. CEP: A conformational epitope prediction server. *Nuc. Acid Res.*, 33: 168-171. <http://bioinfo.ernet.in/publications/cep2k5.pdf>.
16. Zhang, G.L., A.M. Khan, K.N. Guex, N. and M.C. Peitsch, 1997. SWISS-MODEL and the swiss-pdbviewer: An environment for comparative protein modelling. *Electrophoresis*, 18: 2714-2723. <http://www.ncbi.nlm.nih.gov/pubmed/9504803>.
17. Michael, W., R. Andrew and H. Edward, 1999. Widespread intra-serotype recombination in natural populations of dengue virus. *Proc. Natl. Acad. Sci.*, 96: 7352-7357. <http://www.pnas.org/content/96/13/7352.abstract>.

18. Schwede, T., J. Kopp, N. Guex and M.C. Peitsch, 2003. SWISS-MODEL: An automated protein homology-modeling server. *Nucleic Acids Res.*, 31: 3381-3385. <http://www.ncbi.nlm.nih.gov/pubmed/12824332>.
19. Yao, L.H. Y.T. Ming, L.H. Sheng, L.Y. Shin and C.S. Huna, 2001. Immuno pathogenesis of dengue virus infection. *J. Biomed. Sci.*, 8: 377-388. <http://www.ncbi.nlm.nih.gov/pubmed/11549879>.
20. Srinivasan, J.T. Angust and V. Brusica, 2005. MULTIPRED: A computational system for prediction of promiscuous HLA binding peptides. *Nuc. Acids Res.*, 33: 172-179. <http://www.ncbi.nlm.nih.gov/pubmed/15980449>.
21. Modis, Y., S. Ogata, D. Clements and S.C. Harrison, 2005. Variable surface epitopes in the crystal structure of dengue virus type 3 envelope glycoprotein. *J. Virol.*, 79: 1223-1231. DOI: 10.1128/JVI.79.2.1223-1231.2005.
22. Ponsook, K., R. Siburi, S. Supasa, N. Panyadee, A. Songjaeng, A. Jairungsri, C. Puttikhunt, W. Kesinrer, P. Malasit and N. Sittisombut, 2004. Alteration of pr-M cleavage and virus export in pr-M junction chimeric dengue viruses. *J. Virol.*, 78: 2367-2381. DOI: 10.1128/JVI.78.5.2367-2381.2004.
23. Willard, L., A. Ranjan, H. Zhang, H. Monzavi, R.F. Boyko, B.D. Sykes and D.S. Wishart, 2003. VADAR: A web server for quantitative evaluation of protein structure quality. *Nuc. Acids Res.*, 31: 3316-3319. <http://www.ncbi.nlm.nih.gov/pubmed/12824316>.
24. Ho, B.K., A. Thomas and R. Brasseur, 2003. Revisiting the ramachandran plot: Hard-sphere repulsion, electrostatics and H-bonding in the α -helix. *Protein Sci.*, 12: 2508-2522. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2366959>.
25. Westhead, D.R., J.H. Parish and R.M. Twyman, 2002. *Instant Notes Bioinformatics*. 1st Edn., BIOS Scientific Publication, Oxford, ISBN: 10: 1859962726, pp: 304.