e-ISSN: 2248-9126
Vol5|Issue 4| 2015|205-211. Print ISSN: 2248-9118

Indian Journal of Pharmaceutical Science & Research

www.ijpsrjournal.com

HOMOLOGY MODELLING OF NS4B PROTEIN OF DENGUE USING SWISS-PDB VIEWER

Anubrata Paul*, ArpanaVibhuti, V. Samuel Raj, Jayashree Shaw

SRM University, Haryana.Department of Biotechnology.Centre for Drug Design Discovery & Development (CD4), Sonepat – 131 029, Haryana, India.

ABSTRACT

Homology modelling is the method in which 3D model of target which is generally protein is generated from their sequences using a homologous protein which acts as template. The small protein NS4B is a hydrophobic protein, have not been studied extensively. Multiple sequence alignment program was used and identified the corresponding position amino acids in the query sequence of dengue viral protein and templates. A homology modelling method was used for the prediction of the 3D structure of Dengue virus. Modelled structure provided substantial parameters under Ramachandran plot and stereochemical aspects of main chain and side chains through Swiss-Pdb Viewer. Finally the energy minimization protocols scores -3212.67 , -28037.34, -7395.48, -29675.27 KJ/mol for all serotypes of the dengue virus after homology NS4B protein modelling.

Keywords: Dengue Virus, NS4B protein. Homology modeling, Swiss-Pdb viewer, Ramachandran plot, Energy Minimization.

INTRODUCTION

Dengue virus [1] (DENV) is arthropod-borne virus (arbovirus) in the genus Flavivirus (family Flaviviridae) with positive polarity, single – stranded RNA. The RNA is approximately 10.1 kb and is translated into three structural proteins: core protein (C), membrane – associated protein (M) produced as a precursor protein (prM) and envelop protein (E). Additionally, there are 7 non-structural proteins (NS), including NS1, NS2a, NS2b, NS3, NS4a, NS4b, NS5. All DENV group into four genetically related but antigenically distinct serotypes (DENV-1, -2, -3 and -4) within the dengue (DEN) antigenic complex. Dengue is primarily transmitted by *Aedesaegypti* and *Aedesalbopictus* mosquito. It is prevalent in tropic and sub-tropic [2] areas where the vector resides. Approximately 2.5 billion people are at risk of contractin g dengue in endemic areas, and as many as 50 million people are infected, 500,000 hospitalizations and 20,000 deaths with DENV each year by WHO. Dengue virus was first isolated from Japan in 1942 by Hotta. Dengue affects children in Southeast Asia and is characterized by increased vascular permeability, plasma leakage, haemorrhagic [3] manifestations and thrombocytopenia. Both DHF and DSS can be fatal and can lead to death among the patients.

The small protein [4] NS4B has not been studied extensively. NS4B is highly hydrophobic membrane protein which appears to have two hydrophobic segments (residues 1 to 56 and residues 56–93) which are probably associated to the ER lumen side of the membrane and supposedly three C-terminal TM segments (residues 93 to 146, residues 146 to 190 and residues 190 to 248). NS4B is capable of interfering with phosphorylation of STAT1 blocking the IFN-α/β induced signal transduction cascade. NS4B is also a negative modulator of the NS3 helicase function, being this modulation dependent on the conformation of NS4B. NS4B might function cooperatively in viral replication and the anti-host response.

Knowledge on the three-dimensional (3D) structure of a protein [23] is crucial towards understanding its function. This technique relies upon the alignment of a protein sequence of unknown structure [13] (target) to a homologue of known structure (template). Homology modelling is routinely used in many applications However,

Corresponding Author:-**Anubrata Paul** Email: paul.anubiotech@hotmail.com

potential problems can occur in structural determination when the target protein and template have less than 25% sequence identity (based on an average domain length of 80 amino acids). Nevertheless, sufficiently long alignments can still infer structural similarity, even when the sequence similarity is below 25%.

MATERIALS AND METHODS

Searching for serotype sequences of DENV

We searched the National Centre for Biotechnology Information (NCBI) website for the NS4B of Dengue serotypes. Total number of sequences found was 1217, in which more than 600 were randomly selected. The sequences format utilized for saving downloaded NS4b sequences in each DENV serotypes were 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 ClustalWsoftware .

Multiple sequence alignment

Protein sequence alignments [24] and comparisons were done with a BLAST (Basic Local Alignment Search Tool) program, BLASTp against database specification of non-redundant protein which were available at the National Centre for Biotechnology Information (NCBI) Web server (http://www.ncbi.nlm.nih.gov/blast/).Amino acid [14] sequences were obtained from NCBI sequence viewer 2.0 available at (http://www.ncbi.nlm.nih.gov). Multiple sequence alignments[12] were done using ClustalW 1.82, available at the European Bioinformatics Institute (EBI) Web server. To predict the conserved secondary structure profiles, a multiple sequence alignment program was used and identified the corresponding position amino acids in the query sequence of dengue viral protein [15] and templates.

Homology modelling using Swiss Model

"Alignment Mode" was selected among the modeling [25, 6] section for modelling of proteins based on templates. FASTA sequences of both target and template was used for the alignment purpose in order to build a homology model. Selection of target and template sequence along with the name of chain of template was provided in the next step. Based on the template 4 different models were obtained. QMEAN 4 and QMEAN- Z score was given for all the models as an output.

QMEAN [22] (described as Qualitative Model Energy Analysis) can be defined as a composite scoring function for both the estimation of the global quality of entire models as well as for the local per-residue analysis of different regions within a model. The global QMEAN scoring function is a linear combination of different structural descriptors. Here the local geometry was analysed by a torsion angle potential over three consecutive amino acids. In this analysis [15] two different distance-dependent interaction potentials were used to assess long-range interactions. The first is a residue-level implementation based on C-beta atoms only and the second an all-atom potential which is able to capture more details of the model [19]. A solvation potential described here states the burial status of the residues. The QMEAN scoring function returned a reliability score for the whole model which was used in order to compare and rank alternative models of the same target of NS4B protein. The quality estimate ranged between 0 and 1 with higher values for better models except in the case of membrane proteins where lower values [14] can be seen. Additionally, the pseudo energies of the four contributing statistical potential terms were provided as well as the agreement values between predicted and measured features from the sequence and model, respectively. The comparison of the differences of the terms among the models [7] helped in understanding the reason for the differences in the estimated model quality. In addition to the raw scores, Z-score of the QMEAN composite score as well as all terms were provided. In this case the Z-score represented an estimate of the absolute quality of the model by relating it to scores obtained for reference structures solved experimentally by X-ray crystallography. The model's QMEAN score was compared to the scores obtained for PDBs of similar size which is model size of +/- 10% and a Z-score was calculated when the scores were normalized to mean 0 and standard deviation 1. In analogy, Z-scores are calculated for all four statistical potential terms as well as the agreement terms included in the QMEAN score.

The analysis of these Z-scores can help identifying the geometrical features responsible for an observed large (negative) QMEAN Z-score. Models of low quality were expected to have strongly negative QMEAN Z-scores which clearly describe that the model's QMEAN score was several standard deviations lower as expected for experimental structures of similar size except in the case of membrane protein where low scores can be found.

Amino acid Sequence alignment

Amino acid [17] sequence alignment of target and template proteins [21] was derived using the Swiss-Pdb Viewer package [6]. Default parameters were applied and the aligned sequences were inspected and adjusted manually to minimize the number of gaps and insertions. The aligned residues could be seen on the Alignment window in SPDBV. Insertions in the target are proportional to the template were respond as loops with known structure. The estimation of main-chain structures of the loops were modelled by using an efficient and accurate Monte Carlo [18] approach and once the main-chain structure was modelled[10], the side chain atom were affiliated from the Control Panel in SPDBV.

Model optimization and evaluation

The protein model of four serotypes of dengue virus are generated by homology modeling [15] often produce unfavourable bond lengths, bond angles, torsion angles and residues, Non bonded, improper, electrostatic. Side-chain, torsion angles were determined to equal also it is more close to those in the template structure. Therefore, it was essential to minimize the energy to regularize local bond and angle geometry as well as to remove bad contacts. Energy minimisation was applied with the GROMOS96 [18] force field by implementation of Swiss-Pdb Viewer to check for energy [11] criteria in comparison with the potential of mean force derived from a large set of known protein structure. Backbone conformation was analysed by the inspection of Ramachandran plot obtained from Swiss-Pdb viewer [17].

RESULTS & DISCUSSION

CLUSTAL 2.1 Multiple Sequence Alignments

Sequence type explicitly set to Protein
Sequence format is Pearson
Sequence 1: gi|221251|dbj|BAA00395.1| 790 aa
Sequence 2: gi|118505369|gb|ABL01526.1| 643 aa
Sequence 3: gi|573974381|gb|AHG23216.1| 1506aa
Sequence 4: gi|53653745|gb|AAU89376.1| 3387 aa
Start of Pairwise alignments
Aligning...
Sequences (1:2) Aligned. Score: 11.042
Sequences (1:3) Aligned. Score: 11.6456
Sequences (1:4) Aligned. Score: 63.9241
Sequences (2:3) Aligned. Score: 57.0762
Sequences (2:4) Aligned. Score: 73.4059
Sequences (3:4) Aligned. Score: 67.9947
Guide tree file created: [clustalw.dnd]
There are 3 groups
Start of Multiple Alignment
Aligning...
Group 1: Sequences: 2 Score:4852
Group 2: Sequences: 3 Score:16673
Group 3: Sequences: 4 Score:2900
Alignment Score 11279
CLUSTAL-Alignment file created [clustalw.aln]

gi|221251|dbj|BAA00395.1|
NPAVLRKLCIEAKISNTTTDSRCPTQGEATLVEEQDT
NFVCRRTFVDRGW
gi|118505369|gb|ABL01526.1|
LLTLLATVTGGIFLFLMSGRGIGKMTLGMCCIITASIL
LWYAQIQPHWIA
gi|573974381|gb|AHG23216.1|
LLGLMILLTGGAMLFLISGKGIGKTSIGLICVIASSGM
LWMAEIPLQWIA
gi|53653745|gb|AAU89376.1|
LVALLGAMTAGIFLFFMQGKGIGKLSVGLIAIAVAS
GLLWVAEIQPQWIA
gi|221251|dbj|BAA00395.1|
GNGCGLFGKGSLITCAKFKCVTKLEGKIVQYENLKY
SVIVTVHTGDQHQV
gi|118505369|gb|ABL01526.1|
ASIILEFFLIVLLIPEPEKQRTPQDNQLTYVVIAILTVV
AATMANEMGFL
gi|573974381|gb|AHG23216.1|
SAIVLEFFMMVLLIPEPEKQRTPQDNQLAYVVIGILT
LAAIIAANEMGLL
gi|53653745|gb|AAU89376.1|
ASIILEFFLMVLLIPEPEKQRTPQDNQLIYVILAILTIIG
LVAANEMGLI
gi|221251|dbj|BAA00395.1|
GNETTEHGTIATITPQAPTSEIQLTDYGALTLDCSPRT
GLDFNRMVLLTM
gi|118505369|gb|ABL01526.1|
EKTKKDLGLGSITTQQPESNILDIDLRPASAWTLYAV
ATTFVTPMLRHSI
gi|573974381|gb|AHG23216.1|
ETTKRDLGMSKEPGVVSPTSYLDVDLHPASAWTLY
AVATTVITPMLRHTI
gi|53653745|gb|AAU89376.1| EKTKADFGFYQVKT-
--ETTILDVDLRPASAWTLYAVATTILTPMLRHTI
gi|221251|dbj|BAA00395.1|
EKKSWLVHKQWFLDLPLPWTSGASTSQETWNRQDL
LVTFKTAHAKKQEVV
gi|118505369|gb|ABL01526.1|
ENSSVNVSLTAIANQATVLMG--
LGKGWPLSKMDIGVPLLAIGCYSQVNP
gi|573974381|gb|AHG23216.1|
ENSTANVSLAAIANQAVVLMG--
LDKGWPISKMDLGVPLLALGCYSQVNP
gi|53653745|gb|AAU89376.1|
ENTSANLSLAAIANQAAVLMG--
LGKGWPLHRMDLGVPLLAMGCYSQVNP
gi|221251|dbj|BAA00395.1|
VLGSQEGAMHTALTGATEIQTSGTTTIFAGHLKCRL
KMDKLTLKGMSYVM
gi|118505369|gb|ABL01526.1|
ITLTAALFLLVAHYAIIGPGLQAKATREAQKRAAAGI
MKNPTVGRITVID
gi|573974381|gb|AHG23216.1|
LTLTAAVLLLITHYAIIGPGLQAKATREAQKRTAAGI
MKNPTVDGIMTID
gi|53653745|gb|AAU89376.1|
TTLTASLVMLLVHYAIIGPGLQAKATREAQKRTAAG
IMKNPTVDGITVID

Homology modelling with Swiss-Pdb Viewer

When the fig was studied for QMEAN4 score for all 4 models of dengue viral protein it was found that all the four models give a score of approximately 0.1 which was found to be very low. The possible reason given by the Swiss model output was that for membrane protein such a low score is expected. As it is clearly known that NS4B is a type of Viral protein which helps in the export of drugs from membrane so it is quite clear that due to this reason very low score was found. Moreover when QMEAN Z-score was studied for all models it was also found to be

low due to the above reason. In order to further verify the model quality, PSVS was performed for all 4 models so as to choose the best homology model for further studies.

3D structure with Swiss-Pdb Viewer

The quality of individual models can vary significantly from the average accuracy expected for a given target–template similarity or modelling method. SWISS-MODEL workspace provided the graphical plots of Anolea mean force potential, GROMOS empirical force field energy, Verify3D profile evaluation reports were generated to estimate the quality of protein models and template structures.

Ramachandran Plot analysis

Backbone conformation was evaluated by the inspection of the Psi/Phi Ramachandran plot obtained from PROCHECK analysis.TheRamachandran plot of phi/psi distribution in the model is developed using PROCHECK for checking non-GLY residues at the disallowed regions. Standard bond lengths and bond angles of the model were determined using Swiss-Pdb Viewer.

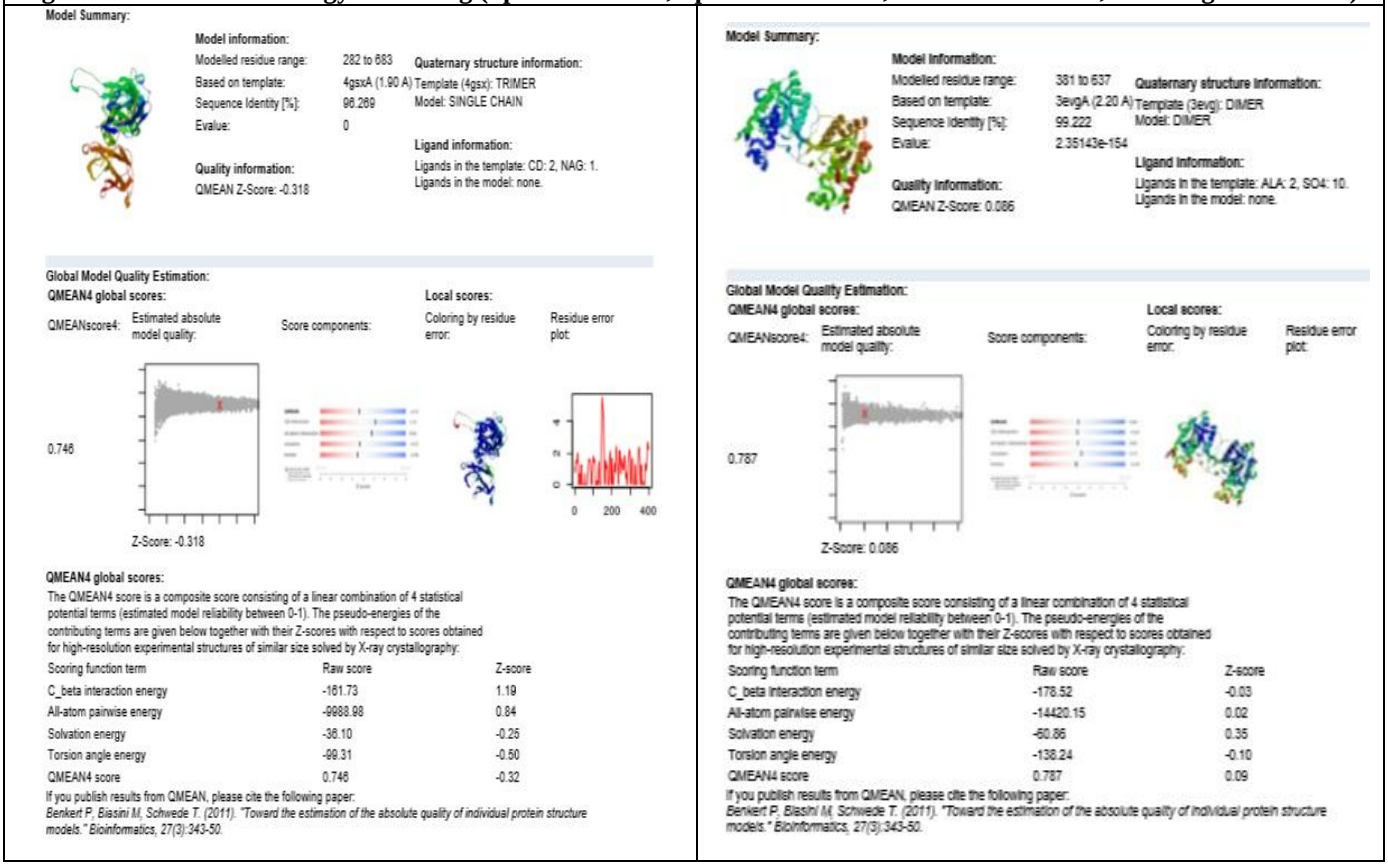
Energy minimization and analysis

The initial models were further optimization by energy minimization using Swiss-Pdb Viewer.

Table 1. Results of Energy Minimization of NS4B protein (DENV-1,2,3,4)

DENV	Residue (KJ/mol)	Bonds (KJ/mol)	Angles (KJ/mol)	Torson (KJ/mol)	Improper (KJ/mol)	Non Bonded (KJ/mol)	Electrostatic (KJ/mol)	Total(E) (KJ/mol)
1	HHTC578-OXTC675	58.84	276.51	431.25	50.61	-2851.10	-1178.80	-3212.67
2	HHTA381-OXTB637	346.43	1950.39	2056.58	376.83	-16521.94	-16245.64	-28037.34
3	HHTB374-OXTB539	91.76	485.57	780.25	99.98	-4844.11	-4008.95	-7395.48
4	HHTA1493-OXTA2092	505.33	2217.76	3124.87	497.99	-19133.20	-16888.64	-29675.27

Fig 1. Results after homology Modelling (up left-DENV 1, up down-DENV 2,down left-DENV 3, down right- DENV 4)



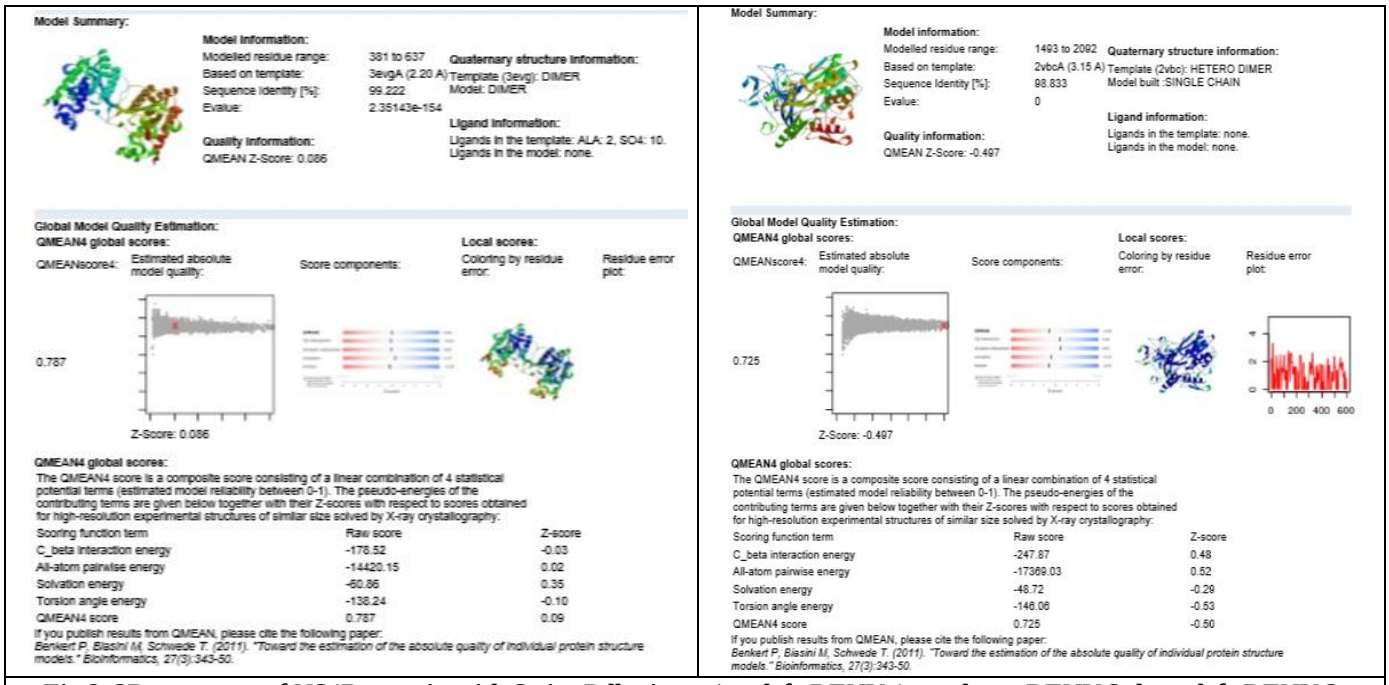


Fig 2. 3D structure of NS4B protein with Swiss-Pdb viewer(up left-DENV 1, up down-DENV 2,down left-DENV 3, down right- DENV 4)

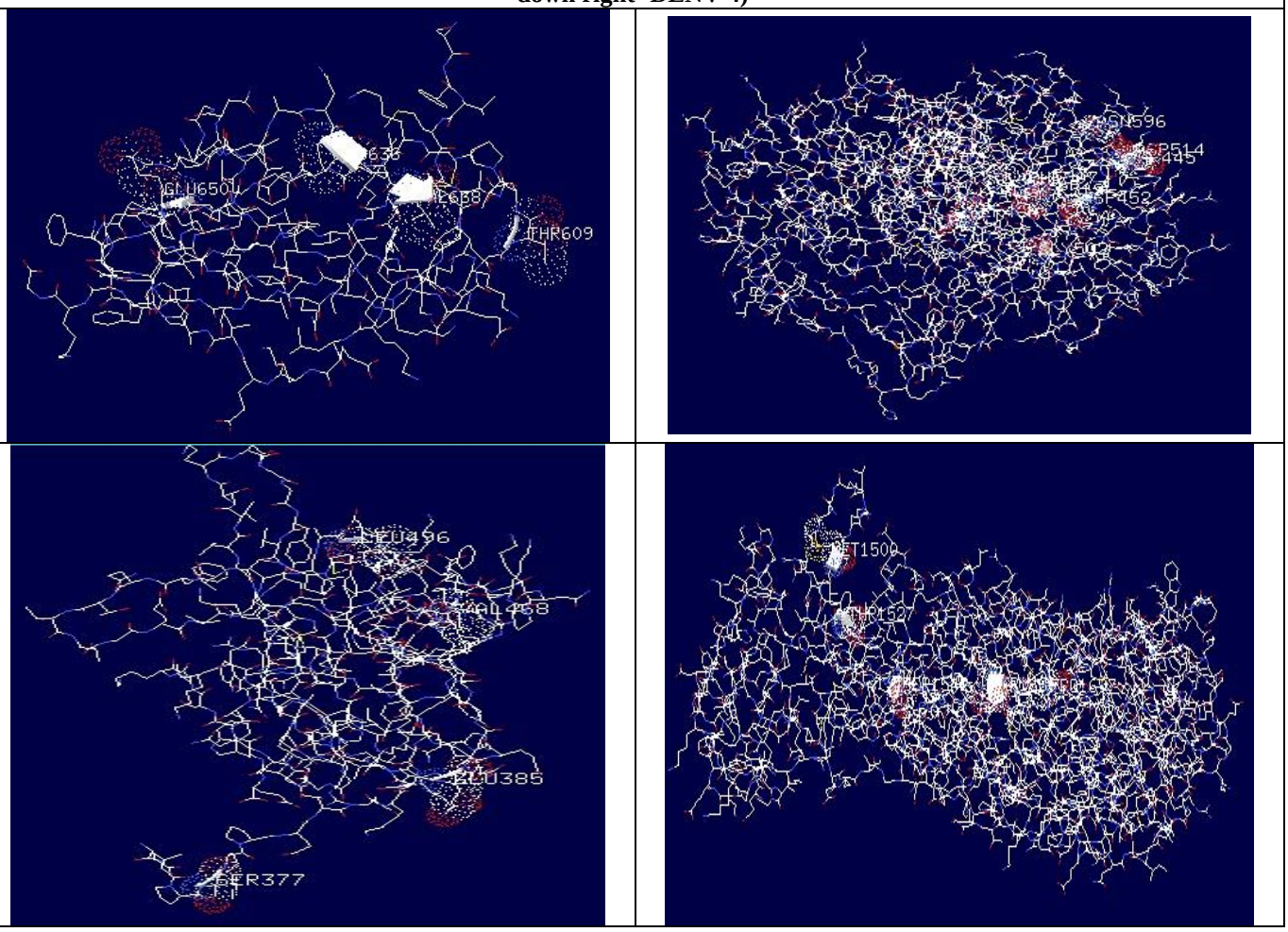
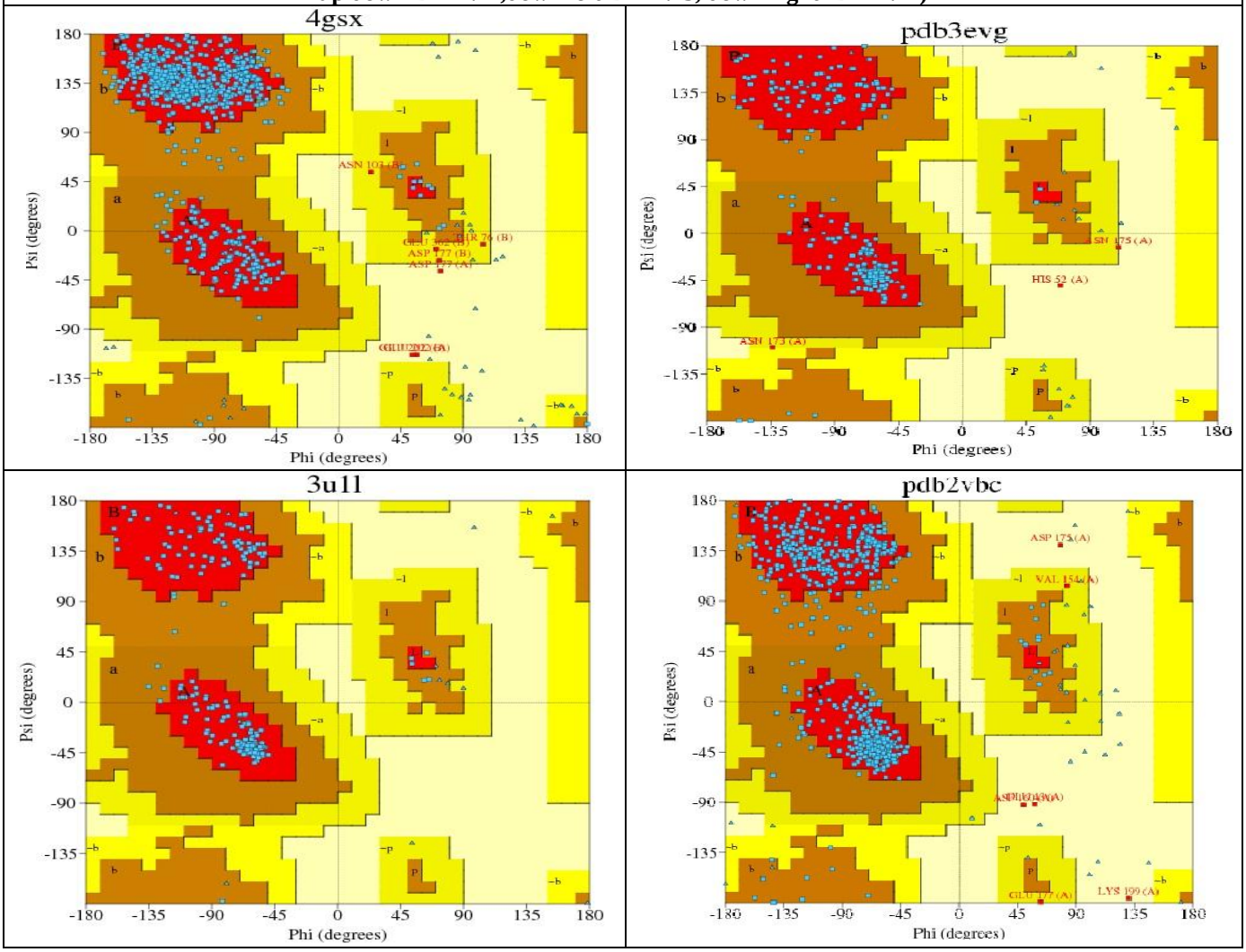


Fig 3. Graphical Representation of Ramachandran Plot of NS4B protein after homology modelling(up left-DENV 1, up down-DENV 2,down left-DENV 3, down right- DENV 4)



CONCLUSION

NS4B is a dengue viral protein both whose functions are still not clear. It is suggested that NS4B may be involved in viral replication. In the current study, a novel approach to conventional homology modelling has been developed. The approach takes advantage of the 3D conformational space corresponding to the template's shape and size characteristics as well as the existence of ligands and substrates, which are used to restrain the folding of the target (query) protein. The ultimate aim is to judge whether homology modelling approach did make a significant difference and improvement to conventionally built model. Homology models of NS4B protein using different servers and software it can be concluded that these homology models can be used for the development of the different inhibitors or drugs by using these homology models for docking studies with different inhibitors of NS4B protein to check the further authencity of the homology model structures. Moreover it can be concluded that these servers and softwares are user friendly and they can be used to generate homology models of nonstructural protein of dengue.

REFERENCES

1. Ven Lim, MohdBasyaruddin A Rahman. *BMC Bioinformatics*, 12(13), 2011, 24.
2. Monath TP. Dengue: the risk to developed and developing countries. *Proc. Natl. Acad. Sci. USA*, 91, 2004, 2395–2400.
3. Erum K, Mehreen K. Demographic and Clinical Features of Dengue Fever in Pakistan from 2003–2007: A Retrospective Cross-Sectional Study. *PLoS ONE*, 5(9), 2010, 12505.
4. Rubing C and Nikos V. *Viruses*, 2011, 3, 1562-1608
5. Leopoldo G, Claudia V. *Viruses*, 3, 2011, 1739-1756.

6. Arnold K, Bordoli L, Kopp J, Schwede T. The SWISS-MODEL Workspace: A web-based environment for protein structure homology modelling. *Bioinformatics*, 22, 2006, 195-201.
7. Westbrook J, *et al*. *Nucleic Acids Res*, 31, 2003, 489– 491.
8. Bairoch A. *Nucleic acid Res*, 33, 2005, D154-D159
9. Tramontano A, Leplae L, Morea V. *Proteins*, 5, 2001, 22–38.
10. Marti-Renom MA. Biophysics. *Biomol.structure*, 29, 2009, 291-325.
11. Kopp J, Schwede T. *Pharmacogenomic*, 5, 2004, 405-416.
12. Hillisch A. *Drug discovery today*, 9, 2004, 659-669.
13. Shewchuk LM, *et al*. Structure of the Tie2 RTK domain. 8, 2000, 1105-13.
14. Arnold K, *et al*. Schwede. *Bioinformatics*, 22(2), 2006, 195-201
15. Sali A, *et al*. *J. Mol. Biology*, 234, 1993, 779–815.
16. Sutcliffe MJ, *et al*. *Protein Eng*, 1, 1987, 377–384.
17. Raj KP, *et al*. *J. Chem. Pharm*. Res., 2(2), 2010, 440-451.
18. Gunsteren WFV. iomolecular Simulations: The GROMOS96 Manual and User Guide. 1996.
19. Melo EF. *J. Mol. Biol*, 277, 1998, 1141–1152.
20. Altschul SF, *et al*., *J. Mol. Biol*., 215, 1990, 403–410.
21. Altschul AF. *Nucleic Acids Res*, 25, 1997, 3389–3402.
22. Huang X. *Adv. Appl. Math*, 12, 1997, 337–357.
23. Lovell SC *et al*., Proteins, 40, 2000, 389–408.
24. Kumar R *et al*. *International Journal of Biological & Pharmaceutical Research*, 5(4), 2014, 354-363.
25. Dhananjeyan K. *Journal of Molecular Modeling*, 15(5), 2009, 507.