Chapter-3

Non-Structural (NS) proteins sequence alignment of viruses (DENV, WNV,YFV) and its Molecular modelling

The non-structural protein sequences of three viruses (DENV, WNV, and YFV) were aligned for knowing its relationship to each other in CLUSTAL W alignment online tool. The non-structural proteins in three viruses are playing respective functions; they include non-structural protein NS1 (NS1), non-structural protein NS2A (NS2A), non-structural protein NS2B (NS2B), non-structural protein NS3 (NS3), non-structural protein NS4A (NS 4A), non-structural protein NS4B (NS4B) and non-structural protein NS5 (NS5).In this chapter we are interested to point out the relation between common protein sequences (non-structural proteins) of three viruses' homology using protein sequence alignment and its molecular modelling, validation through Ramachandran plot.

CLUSTAL W, an improvement on the original CLUSTAL program was introduced in 1994 (Thompson, et. al., 1994), and its relative speed and sensitivity soon made it the method of choice for biologists. The enhanced sensitivity was due to three improvements incorporated in to CLUSTAL W: the use of a weighted sum-of-pairs, the use of varying gap penalties and the use of neighbour-joining instead of UPGMA in the generation of the phylogenetic.

In older alignment algorithms, single-weight matrices were generated from the pair wise alignments. A single-weight matrix assumes that the sequences in a group to be aligned are all equally divergent from each other. As long as this assumption is true then the use of single weight matrices is justified. However, if the sequences in an alignment group are too similar, then a single-weight matrix introduces a bias in favour of redundant sequences. If the sequences are divergent, then mismatches in sequence become more prevalent but single-weight matrices do not account for them. CLUSTAL W assigns individual weights to sequences, in part by using neighbour joining to construct the phylogenetic tree: weights are assigned according to the tree branch length, which is a measure of their evolutionary distance. This means that similar or 4 redundant sequences will be down weighted while more divergent sequences are up weighted (Chenna, et al.,2003).

We used to align the protein sequences for their homology relationship in Clustal W tool. To making of homology alignment in between three viruses are downloaded from the NCBI database. The accession number for the viruses NS1 in NCBI database was denoted for Dengue Virus NP_722461.1, West Nile Virus NP_776015.1 and Yellow Fever Virus

NP_776002.1. The NS1 protein of three viruses sequence length was 352 amino acids. The NS1 sequences alignment in CLUSTAL W showed that the Dengue Virus NS1 relationship score with the West Nile Virus is 42.05 and same protein relation with the Yellow Fever Virus was 51.42 whereas the West Nile Virus has 44.03 relational score value with the Yellow Fever Virus (figure 15). Though the sequence of three NS1 proteins from three viruses was 352, the identity-similarity score is different from each other.

The NS2A protein sequences of three virus NCBI accession numbers were NP_733808.1 for Dengue, NP_776016.1 for West Nile and NP_776003.1 for Yellow Fever Virus. The sequence length NS2 of Dengue was 218 amino acids; West Nile Virus was 231 amino acids and 224 amino acids sequence length for Yellow Fever Virus. The sequence alignment for these proteins revealed their relationship score. The relationship score in between Dengue NS2A and NS2A of West Nile Virus was 22.94, the score between Dengue and Yellow Fever Virus was 19.72 and 26.34 relationship score was showed between West Nile and Yellow Fever Virus (figure 16).

NS2B protein sequence accession number NP_733809.1 for Dengue Virus, NP_776017.1 for West Nile Virus and NP_776004.1 accession number for Yellow Fever Virus were used to download from NCBI database. The CLUSTAL W alignment of three sequences was showed their relationship. The Dengue Virus NS2B protein has relation with West Nile Virus was 29.23. The Dengue Virus has 30.77 relational score with the Yellow Fever Virus. The West Nile Virus has the 27.69 relational score with the Yellow Fever Virus (figure 17).

The Dengue Virus NS3 protein sequence accession number in NCBI database was NP_722463.1, West Nile NS3 was NP_776018.1and Yellow Fever NS3 was NP_776005.1. The Dengue, West Nile, Yellow Fever Viruses'NS3 protein sequence alignment in CLUSTAL W revealed their relation. The Dengue Virus NS3 showed 62.36 relational score with the NS3 protein of West Nile Virus and 51.86 relational score with the Yellow Fever Virus. The West Nile virus has showed 52.99 relation score with Yellow Fever Virus (figure 18).

The Dengue Virus NS4A protein sequence in NCBI database allocated accession number was NP_733810.1; West Nile Virus NS4A protein accession number was NP_776019.1 whereas the allocated accession number for Yellow Fever Virus NS4A was

NP_776006.1.The relationship in between NS4A of Dengue, West Nile and Yellow Fever Virus was studies in CLUSTAL W sequence alignment tool. The 36.15 relational score showed Dengue Virus with the West Nile Virus whereas 30.16 score with the Yellow Fever Virus. The West Nile Virus showed 30.16 relational score with The Yellow Fever Virus (figure 19).

The NCBI allocated accession number for NS4B protein sequence of Dengue Virus was NP_733811.1; NS4B protein sequence of West Nile Virus was NP_776021.1 whereas the NS4B protein sequence of Yellow Fever Virus was NP_ 776008. The evolutionary relationship among Dengue, West Nile and Yellow Fever Virus'sNS4B protein was revealed that the Dengue Virus has 38.55 relational score with West Nile Virus and 34.94 score with Yellow Fever Virus. The West Nile Virus has the 30.4 relational score with the Yellow Fever Virus (figure 20).

The accession number of the Dengue Virus NS5 protein in NCBI data base was NP_722465.1, West Nile Virus NS5 protein accession number was NP_776022.1 and Yellow Fever virus accession number was NP_776009.1. The evolutionary relationship for NS5 protein of Dengue Virus with the West Nile Virus was 65.18 similarity-identity score whereas 60.29 similarity-identity score with Yellow Fever Virus. The West Nile Virus has the 60.44 similarity-identity score with the Yellow Fever Virus (figure 21).

The template crystallography structures against PDB database in BLAST P program. Protein database is the most important protein resource at NCBI. It maintains the text record for individual protein sequences, derived from many different resources such as NCBI Reference Sequence (RefSeq) project, Genbank, PDB and UniProtKB/SWISS-Prot. Protein records are present in different formats including FASTA and XML and are linked to other NCBI resources. It also provides the pre-determined sets of similar and identical proteins for each sequence as computed by the BLAST. The Structure database of NCBI contains 3D coordinate sets for experimentally-determined structures in PDB that are imported by NCBI. There is another database in protein known as Protein Clusters database which contains sets of proteins sequences that are clustered according to the maximum alignments between the individual sequences as calculated by BLAST (Sayers, E., 2013).

Sequence similarity searching to identify homologous sequences is one of the first and most informative steps in any analysis of newly determined sequences. Modern protein sequence databases are very comprehensive, so that more than 80% of metagenomic sequence samples typically share significant similarity with proteins in sequence databases. Widely used similarity searching programs, like BLAST (Altschul, et al., 1997), PSI-BLAST (Altschul, et al., 1997), SSEARCH (Smith and Waterman, 1981); Pearson, 1991), FASTA (Pearson and Lipman, 1988) and the HMMER3 (Johnson, et al., 2010) programs produce accurate statistical estimates, ensuring protein sequences that share significant similarity also have similar structures. Similarity searching is effective and reliable because sequences that share significant similarity can be inferred to be homologous; they share a common ancestor.

The concept of homology is a common evolutionary ancestry, is central to computational analyses of protein and DNA sequences. Homology in between two sequences or structures share more similarity than would be expected by chance; when excess similarity is observed, the simplest explanation for that excess is that the two sequences did not arise independently, they arose from a common ancestor. Common ancestry explains excess similarity (other explanations require similar structures to arise independently); thus excess similarity implies common ancestry.

Homology (common ancestry and similar structure) can be reliably inferred from statistically significant similarity in a BLAST, but to infer that two proteins are homologous does not guarantee that every part of one protein has a homolog in the other. BLAST, calculate local sequence alignments; local alignments identify the most similar region between two sequences. For single domain proteins, the end of the alignment may coincide with the ends of the proteins, but for domains that are found in different sequence contexts in different proteins, the alignment should be limited to the homologous domain, since the domain homology is providing the sequence similarity captured in the score. When local alignments end within a protein, the ends of the alignment can depend on the scoring matrix used to calculate the score. In particular, scoring matrices like BLOSUM62, which is used by BLASTP, is designed to detect very distant similarities, and have relatively low penalties for mismatched residues. As a result, a homologous region that is 50% identical or more can

be extended outside the homologous domain into neighbouring non-homologous regions. This is a common cause of errors with iterative methods like PSI-BLAST (Gonzalez and Pearson, 2010), but can be reduced by limiting extension in later iterations (Li, et al., 2012).

The non-structural protein sequences of three viruses from NCBI database were used to search its respective template protein structures in BLAST tool against protein data bank. The template structures are protein x ray crystallographic structures have similarity in the query sequence. The protein sequence aligned with the sequences of protein databank contained crystal structures sequences. Based on the alignment score the best fit template for query sequence was selected. Based on the template structure the protein sequence of query builds three dimensional structures.

The template structures identified for NS1 protein of Dengue Virus in BLAST was "Chain A, Dengue Type2 Virus Non-structural Protein 1 (NS1) form 1 Crystal" and its protein databank id is 4O6B. The query sequence NS1 of Dengue showed the 73% identity. West Nile Virus NS1 protein search for protein data bank generated a template structure named "West Nile Virus Non-structural Protein 1 (NS1) form 2 Crystal with query coverage 100 % and 91 % identity. The Yellow Fever Virus NS1 protein template identified in BLAST search was "Chain A, West Nile Virus Non-structural Protein 1 (NS1) Form 2 Crystal" with PDB id 4O6C. The sequence identity was 43% and 100 % query coverage.

The NS2A protein of Dengue against protein databank in BLAST search produced a template structure "Chain A, Solution Structure of The Trans-Membrane Domain of the NS2A of Dengue" with query coverage 12% and identity 64%. West Nile virus template identified in BLAST search was "Chain A, Structure Analysis of The Global Metabolic Regulator Crc from Pseudomonas Aeruginosa" with query coverage 22% and identity 31%. The protein structure "ChainA, Crystal Structure of The Closed form of Pseudomonas Aeruginosa Spm-1" was template for Yellow Fever Virus NS2A sequence with query coverage 22%, and identity 32%.

Three Viruses (DENV, WNV and YFV) NS2B protein template structure identification in BLAST generated the structures. "Chain A, Crystal Structure of Dengue Virus1 NS2B-NS3 Protease Active Site Mutant" with query coverage 36%, and 100% identity for Dengue Virus."ChainA, West Nile Virus NS2B/NS3 Protease in Complex with

Bz-nle-lys-arg-arg-h with query coverage 36% and identity 98% for WestNile virus. "Chain A, Crystal Structure of the West Nile Virus NS2B-NS3 Protease Complexed with Bovine Pancreatic Trypsin Inhibitor" with 32% query coverage and 43 % identity.

The Dengue Virus NS3 protein templates PDB structure was "Chain A, Insertion Mutant E173gp174 of the Ns3 Protease-Helicase from Dengue Virus". It showed 100 % query coverage and 73% identity with the NS3 sequence of Dengue. The West Nile Virus NS3 protein template PDB structure was "Chain A, Crystal Structure of the Ns3 Protease-Helicase from Murray Valley Encephalitis Virus"with 99% query coverage and 81% identity. The Yellow Fever virus NS3 protein template structure was "Chain A, Crystal Structure of Yellow Fever Virus NS3 Helicase" with 70% query coverage and 100 % identity.

The Dengue Virus NS4A template structure was "Chain A, Crystal Structure of A Putative Cyclic Nucleotide-Binding Protein (Gmet_1532) from Geobacter Metallireducens Gs-15 At 1.90 A Resolution" with query coverage 35% and identity 30 %. The West Nile Virus NS4A template structure was "Chain A, The Structure of The Processed form of Threonine Deaminase Isoform 2 from Solanum Lycopersicum" with 22% query coverage and 54% identity. The Yellow Fever Virus NS4A template structure was "Chain A, RuvaFrom Mycobacterium Tuberculosis" with 28 % query coverage and 42% identity.

The Dengue virus NS4B protein sequence search for template in BLAST generates the "chain A, Elongation Factor G1 (Pseudomonas Aeruginosa) in Complex with ArgyrinB Protein Structure". It has 13 % query coverage and 35% identify with the query NS4B sequence. The West Nile Virus NS4B sequence search for template in BLAST not generated any template so this query was modelled build using threading method. The Yellow Fever Virus NS4B template search in BLAST was produced "Chain A, Eris Sting in Complexed with Ligand" structure with 28% query coverage and 25% identity.

The template for the NS5 protein of the Dengue Virus found in BLAST search was " ChainA, Dengue Virus Full Length Ns5 Complexed with Sah" with query coverage 98 % and 81% identity. The template search for West Nile Virus NS5 protein founded in BLAST search was "ChainA, Crystal Structure of the Full-Length Japanese Encephalitis Virus NS5" with 100% query coverage and 81% identity. The template for the NS5 of the Yellow Fever

Virus in BLAST search found that "Chain A, Dengue Virus Full Length NS5 Complexed with Sah" with 98%query coverage and 60% identity.

The templates related to query sequences of three viruses' non-structural proteins of Dengue Virus, West Nile and Yellow Fever viruses are together used for the building of three dimensional structures of non-structural proteins. The three dimensional protein structures of non-structural proteins were built in modeller 9 version 10. Modeller generates modelled structures about query, are selected based on DOPE score values. All non-structural proteins of three viruses are modelled and validated through Ramachandran plot. The modelled structures proteins were represented in figure 22, 23 and 24.

The secondary structures that polypeptides can adopt in proteins are governed by hydrogen bonding interactions between the electronegative carbonyl oxygen atoms and the electropositive amide hydrogen atoms in the backbone chain of the molecule. These hydrogen-bonding interactions can form the framework that stabilizes the secondary structure. Many secondary structures with reasonable hydrogen bonding networks could be proposed. Most of the possible secondary structures are not possible due to limits on the configuration of the backbone of each amino acid residue (Alice, et al.,2011). Understanding these limitations will help you to understand the secondary structures of proteins. This can obtained through Ramachandran plot regions in the set of possible amino acid configurations that are allowed and disallowed.

Modelled proteins were validated in SAVES server Ramachandran plot. Ramachandran plot calculate the amino acid residues in different positions in a plot (percentage) are indicative of their good structure reliability. TheNS1 allowed region amino acid residues of Dengue was 88.9, West Nile 63.2and Yellow Fever was 89.9. The non-structural protein NS2 of Dengue allowed region amino acid percentage was86.7; West Nile Virus was 86.0 and 86.0. Allowed region amino acid residues of NS2B protein in Dengue Virus was 78.1, in West Nile Virus was 88.6 and in Yellow Fever Virus was 92.7. The Dengue Virus NS 3 protein has allowed region amino acid residue percentage was 90.9, West Nile has 89.0 and Yellow Fever virus has 89.3. Ns4a protein in Dengue virus has 82.4 percentages allowed region amino acids, West Nile percentage was 88.6 and Yellow Fever percentage in allowed region was84.0. The allowed region amino acids in Dengue Virus

protein NS4B was 93.0, West Nile NS4B has 85.9 percentages and Yellow Fever Virus has 81.1 percentage amino acids in allowed region. NS 5 protein of Dengue Virus has 93.0 percentage amino acids in allowed region whereas the West Nile Virus has 85.9 percentages and Yellow Fever Virus NS5 protein has 81.1 percent amino acids in allowed region.

```
gi|25014063|ref|NP_722461.1|      DSGCVINWKGRELKCGSGIFVTNEVHTWTEQYKFQADSPKRLSAAIGKAW 50
gi|27735303|ref|NP_776015.1|      DTGCAIDIGRQELRCGSGVFIHNDVEAWMDRYKFYPETPQGLAKIIQKAH 50
gi|27735290|ref|NP_776002.1|      DQGCAINFGKRELKCGDGIFIFRDSDDWLNKYSYYPEDPVKLASIVKASF 50

gi|25014063|ref|NP_722461.1|      EEGVCGIRSATRLENIMWKQISNELNHILLENDMKFTVVVGDVSGILAQG 100
gi|27735303|ref|NP_776015.1|      AEGVCGLRSVSRLEHQMWEAIKDELNTLLKENGVDLSVVVEKQNGMYKAA 100
gi|27735290|ref|NP_776002.1|      EEGKCGLNSVDSLEHEMWRSRADEINAIFEENEVDISVVVQDPKNVYQRG 100

gi|25014063|ref|NP_722461.1|      KKMIRPQPMEHKYSWKSWGKAKIIGADVQNTTFIIDGPNTPECPDNQRAW 150
gi|27735303|ref|NP_776015.1|      PKRLAATTEKLEMGWKAWGKSIIFAPELANNTFVIDGPETEECPTANRAW 150
gi|27735290|ref|NP_776002.1|      THPFSRIRDGLQYGWKTWGKNLVFSPGRKNGSFIIDGKSRKECPFSNRVW 150

gi|25014063|ref|NP_722461.1|      NIWEVEDYGFGIFTTNIWLKLRDSYTQVCDHRLMSAAIKDSKAVHADMGY 200
gi|27735303|ref|NP_776015.1|      NSMEVEDFGFGLTSTRMFLRIREТNTTECDSKIIGTAVKNNMAVHSDLSY 200
gi|27735290|ref|NP_776002.1|      NSFQIEEFGTGVFTTRVYMDAVFEYTIDCDGSILGAAVNGKKSAHGSPTF 200

gi|25014063|ref|NP_722461.1|      WIES-EKNETWKLARASFIEVKTCIWPKSHTLWSNGVLESEMIIPKIYGG 249
gi|27735303|ref|NP_776015.1|      WIES-GLNDTWKLERAVLGEVKSCTWPETHTLWGDGVLESDLIIPITLAG 249
gi|27735290|ref|NP_776002.1|      WMGSHEVNGTWMIHTLEALDYKECEWPLTHTIGTS-VEESEMFMPRSIGG 249

gi|25014063|ref|NP_722461.1|      PISQHNYRPGYFTQTAGPWHLGKLELDFDLCEGTTVVVDEHCGNRGPSLR 299
gi|27735303|ref|NP_776015.1|      PPSNHNRRPGYKTQNQGPWDEGRVEIDFDYCPGTTVTISDSCEHRGPAAR 299
gi|27735290|ref|NP_776002.1|      PVSSHNHIPGYKVQTNGPWMQVPLEVKPEACPGTSVIIDGNCDGRGKSTR 299

gi|25014063|ref|NP_722461.1|      TTTVTGKIIHEWCCRSCTLPPLRFKGEDGCWYGMEIRPVKEKEENLVKSM 349
gi|27735303|ref|NP_776015.1|      TTTESGKLITDWCCRSCTLPPLRFQTENGCWYGMEIRPTRHDEKTLVQSR 349
gi|27735290|ref|NP_776002.1|      STTDSGKVIPEWCCRSCTMPPVSFHGSDGCWYPMEIRPRKTHESHLVRSW 349

gi|25014063|ref|NP_722461.1|      VSA 352
gi|27735303|ref|NP_776015.1|      VNA 352
gi|27735290|ref|NP_776002.1|      VTA 352
```

SeqA	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
1	gi|25014063|ref|NP_722461.1|	352	2	gi|27735290|ref|NP_776002.1|	352	42.05
1	gi|25014063|ref|NP_722461.1|	352	3	gi|27735303|ref|NP_776015.1|	352	51.42
2	gi|27735290|ref|NP_776002.1|	352	3	gi|27735303|ref|NP_776015.1|	352	44.03

Figure 15: The non-structural protein NS1 Sequence alignment of three viruses (DV, WV and YFV) in CLUSTAL W tool. The alignment based on Gonnet PAM 250 matrix. (The symbols like '* ': fully conserved residues, ': ' : conserved and ' . ': semi conservative residue alignments)

375253


```
gi|27735304|ref|NP_776016.1|      YNADMIDPFQLGLMVVFLATQEVLRKRWTAKISIPAIMLALLVLVFGGIT 50
gi|27735291|ref|NP_776003.1|      ---GKIHAVPFGLVSMMIAMEVVLRKRQGPKQMLVGGVVLLGAMLVGQVT 47
gi|25014064|ref|NP_733808.1|      -GSGEVDSFSLGLLCISIMIEEVMRSRWSRKMLMTGTLAVFLLLTMGQLT 49

gi|27735304|ref|NP_776016.1|      YTDVLRYVILVGAAFAEANSGGDVVHLALMATFKIQPVFLVASFLKARWT 100
gi|27735291|ref|NP_776003.1|      LLDLLKLTVAVGLHFHEMNNGGDAMYMALIAAFSIRPGLLIGFGLRTLWS 97
gi|25014064|ref|NP_733808.1|      WNDLIRLCIMVGANASDKMGMG-TTYLALMATFRMRPMFAVGLLFRRLTS 98

gi|27735304|ref|NP_776016.1|      NQESILLMLAAAFFQMAYYDAKNVLSWEVPDVLNSLSVAWMILRAISFTN 150
gi|27735291|ref|NP_776003.1|      PRKRLVLTLGAAMVEIALGGVMG---GLWKYLNAVSLCILTINAVASRK 143
gi|25014064|ref|NP_733808.1|      ---REVLLLTVGLSLVASVELPN---SLEELGDGLAMGIMMLKLLTDFQ 141

gi|27735304|ref|NP_776016.1|      TSNVVVPLLALLTPGLKCLNLDVYRILLLMVGVGSLIKEKRSSAAKKKGA 200
gi|27735291|ref|NP_776003.1|      ASNTILPLMALLTP-VTMAEVRLAAMFFCAVVIIGVLHQNFKDTSMQKTI 192
gi|25014064|ref|NP_733808.1|      SHQLWATLLSLTFVKTTFSLHYAWKTMAMILSIVSLFPLCLSTTSQKTTW 191

gi|27735304|ref|NP_776016.1|      CLICLALAS-TGVFNPMILAAGLMACDPNRKR 231
gi|27735291|ref|NP_776003.1|      PLVALTLTSYLGLTQPFLGLCAFLATRIFGRR 224
gi|25014064|ref|NP_733808.1|      LPVLLGSLGCKPLTMFLITENKIWGRK----- 218
```

1	gi|25014064|ref|NP_733808.1|	218	2	gi|27735304|ref|NP_776016.1|	231	22.94
1	gi|25014064|ref|NP_733808.1|	218	3	gi|27735291|ref|NP_776003.1|	224	19.72
2	gi|27735304|ref|NP_776016.1|	231	3	gi|27735291|ref|NP_776003.1|	224	26.34

Figure 16: The non-structural protein NS2A Sequence alignment of three viruses (DV, WV and YFV) in CLUSTAL W tool. The alignment based on Gonnet PAM 250 matrix. (The symbols like '* ': fully conserved residues, ': ' : conserved and ' . ': semi conservative residue alignments)

RS
595.771
K 836

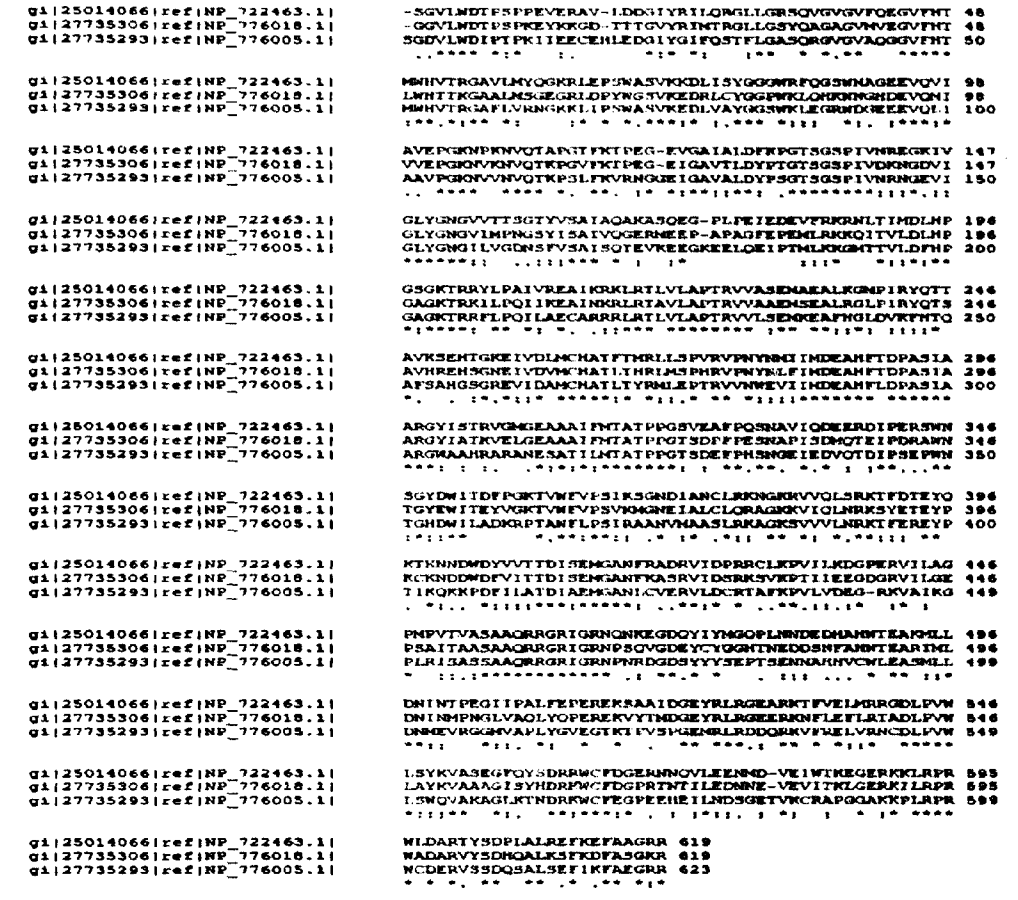
```
gi|25014065|ref|NP_733809.1|      SWPLNEGIMAVGIVSILLS-SLLKNDVPLAGPLIAGGMLIACYVISGSSA 49
gi|27735292|ref|NP_776004.1|      SIPVNEALAAAGLVGVLAG-LAFQEMENFLGPIAVGGLLMMLVSVAGRVD 49
gi|27735305|ref|NP_776017.1|      GWPATEVMTAVGLMFAIVGGLAELDIDSMAIPMTIAGLMFAAFVISGKST 50

gi|25014065|ref|NP_733809.1|      DLSLEKAAEVSWEEEAEHSGASHNILVEVQDDGTMKIKDEERDDTLTILL 99
gi|27735292|ref|NP_776004.1|      GLELKKLGEVSWEEEAEISGSSARYDVALSEQGEFKLLSEEKVPWDQVVM 99
gi|27735305|ref|NP_776017.1|      DMWIERTADITWESDAEITGSSERVDVRLDDDGNFQLMNDPGAPWKIWML 100

gi|25014065|ref|NP_733809.1|      KAILLAISGVYPMSIPATLFVWYFWQKKKQR 130
gi|27735292|ref|NP_776004.1|      TSLALVGAALHPFALLLVLAGWLFHVRGARR 130
gi|27735305|ref|NP_776017.1|      RMACLAISAYTPWAILFSVIGFWITLQYTKR 131
```

SeqA	NameA	LengthA	SeqB	NameB	LengthB	Score
1	gi|25014065|ref|NP_733809.1|	130	2	gi|27735305|ref|NP_776017.1|	131	29.23
1	gi|25014065|ref|NP_733809.1|	130	3	gi|27735292|ref|NP_776004.1|	130	30.77
2	gi|27735305|ref|NP_776017.1|	131	3	gi|27735292|ref|NP_776004.1|	130	27.69

Figure 17: The non-structural protein NS2B Sequence alignment of three viruses (DV, WV and YFV) in CLUSTAL W tool. The alignment based on Gonnet PAM 250 matrix. (The symbols like '*': fully conserved residues, ':' : conserved and '.': semi conservative residue alignments)



1	gi|25014066|ref|NP_722463.1|	619	2	gi|27735306|ref|NP_776018.1|	619	62.36
1	gi|25014066|ref|NP_722463.1|	619	3	gi|27735293|ref|NP_776005.1|	623	51.86
2	gi|27735306|ref|NP_776018.1|	619	3	gi|27735293|ref|NP_776005.1|	623	52.99

Figure 18: The non-structural protein NS3 Sequence alignment of three viruses (DV, WV and YFV) in CLUSTAL W tool. The alignment based on Gonnet PAM 250 matrix. (The symbols like '*, :': fully conserved residues, ':': conserved and '.': semi conservative residue alignments)

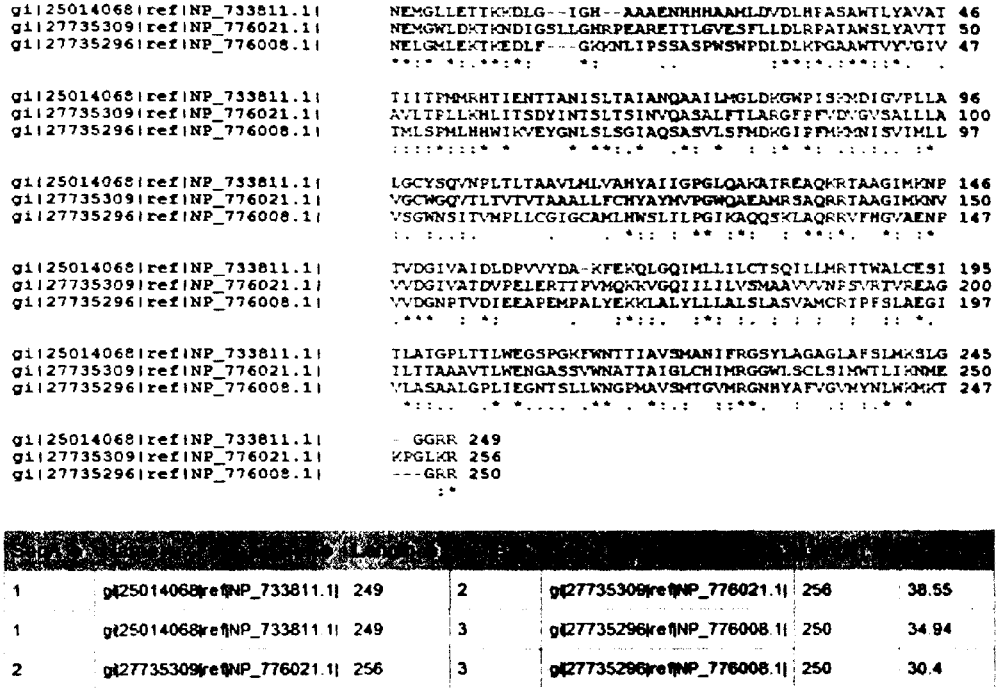
```
gi|25014067|ref|NP_733810.1|     SVSGDLILEIGKLPQHLTQRAQNALDNLVMLHNSEQGGKAYRHAMEELPD 50
gi|27735307|ref|NP_776019.1|     SQIG-LVEVLGRMPEHFMVKTWEALDTMYVVATAEKGGRAHRMALEELPD 49
gi|27735294|ref|NP_776006.1|     -GAAEVLVVLSELPDFLAKKGGEAMDTISVFLHSEEGSRAYRNALSMKPE 49
                                     . ::  :..:*:.:  :  :*:*.: :.  :*:*.:*:* *:. :*:

gi|25014067|ref|NP_733810.1|     TIETLMLLALIAVLTGGVTLFFLSGRGLGKTSIGLLCVIASSALLHMASV 100
gi|27735307|ref|NP_776019.1|     ALQTIVLIALLSVMSLGVFFLLMQRKGIGKIGLGGVILGAATFFCHMAEV 99
gi|27735294|ref|NP_776006.1|     AMTIVMLFILAGLLTSGMVIFFMSPKGISRMSMAMGTMAGCGYLMFLGGV 99
                                  ::  ::*: * .::: *: ::::. :*:.: .:.   : ..  : ::. *

gi|25014067|ref|NP_733810.1|     EPHWIAASIILEFFLMVLLIPEPDRQR 127
gi|27735307|ref|NP_776019.1|     PGTKIAGMLLLSLLLMIVLIPEPEKQR 126
gi|27735294|ref|NP_776006.1|     KPTHISYVMLIFFVLMVVVIPEPGQQR 126
                                     *:  ::: :.**:::**** :**
```

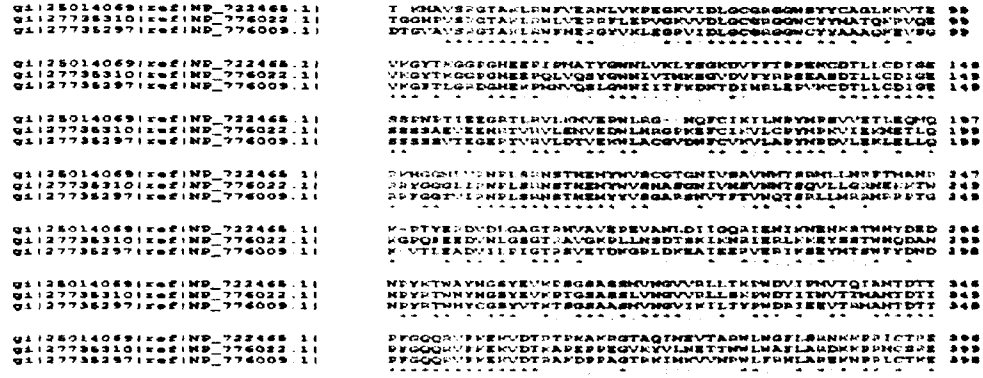
Seq A	Name	Len	Seq B	Name	Len	Score
1	gi|25014067|ref|NP_733810.1|	127	2	gi|27735307|ref|NP_776019.1|	126	36.51
1	gi|25014067|ref|NP_733810.1|	127	3	gi|27735294|ref|NP_776006.1|	126	30.16
2	gi|27735307|ref|NP_776019.1|	126	3	gi|27735294|ref|NP_776006.1|	126	30.16

Figure 19: The non-structural protein NS4A Sequence alignment of three viruses (DV, WV and YFV) in CLUSTAL W tool. The alignment based on Gonnet PAM 250 matrix. (The symbols like '* ': fully conserved residues, ' : ' : conserved and ' . ': semi conservative residue alignments)



1	gi|25014068|ref|NP_733811.1|	249	2	gi|27735309|ref|NP_776021.1|	256	38.55
1	gi|25014068|ref|NP_733811.1|	249	3	gi|27735296|ref|NP_776008.1|	250	34.94
2	gi|27735309|ref|NP_776021.1|	256	3	gi|27735296|ref|NP_776008.1|	250	30.4

Figure 20: The non-structural protein NS4B Sequence alignment of three viruses (DV, WV and YFV) in CLUSTAL W tool. The alignment based on Gonnet PAM 250 matrix. (The symbols like '* *': fully conserved residues, ' : ' : conserved and ' . ': semi conservative residue alignments)



1	gi|25014069|ref|NP_722465.1|	899	2	gi|27735310|ref|NP_776022.1|	905	65.18
1	gi|25014069|ref|NP_722465.1|	899	3	gi|27735297|ref|NP_776009.1|	905	60.29
2	gi|27735310|ref|NP_776022.1|	905	3	gi|27735297|ref|NP_776009.1|	905	60.44

Figure 21: The non-structural protein NS5 Sequence alignment of three viruses (DV, WV and YFV) in CLUSTAL W tool. The alignment based on Gonnet PAM 250 matrix. (The symbols like '* ': fully conserved residues, ' : ' : conserved and ' . ': semi conservative residue alignments)

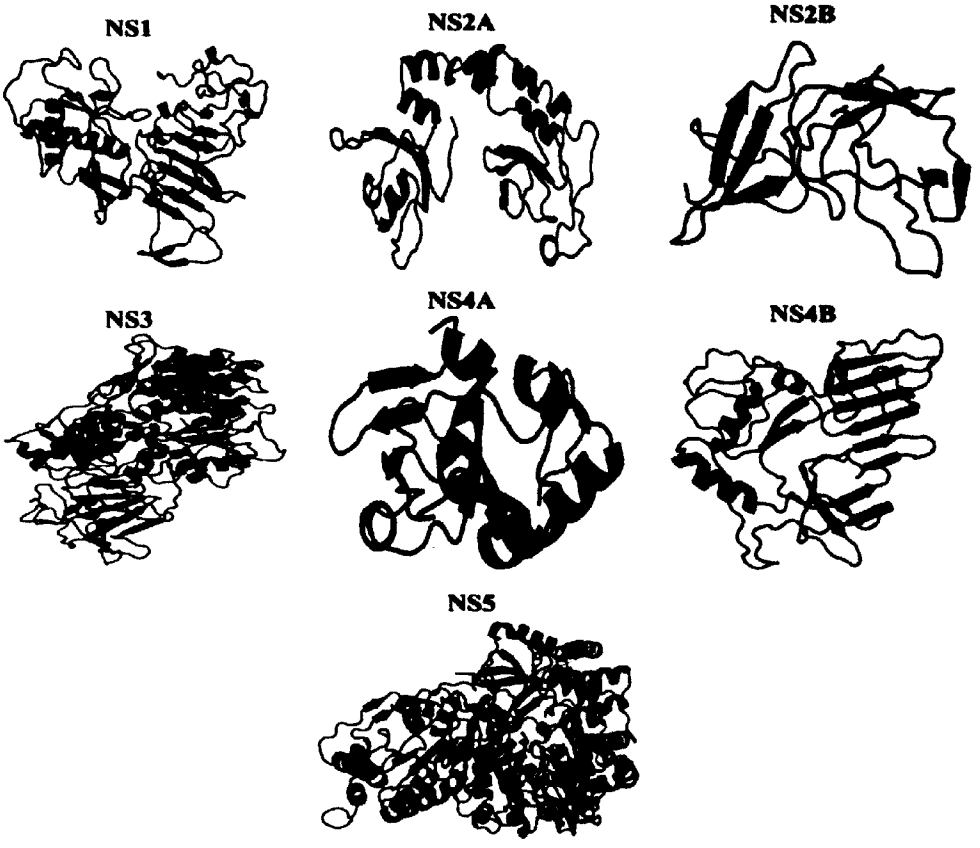


Figure 22: The molecular modelling of the non-structural proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B and NS5) of Dengue Virus. The modelling of proteins done in MODELLER 9.10 and molecular visualization was seen in Pymol.

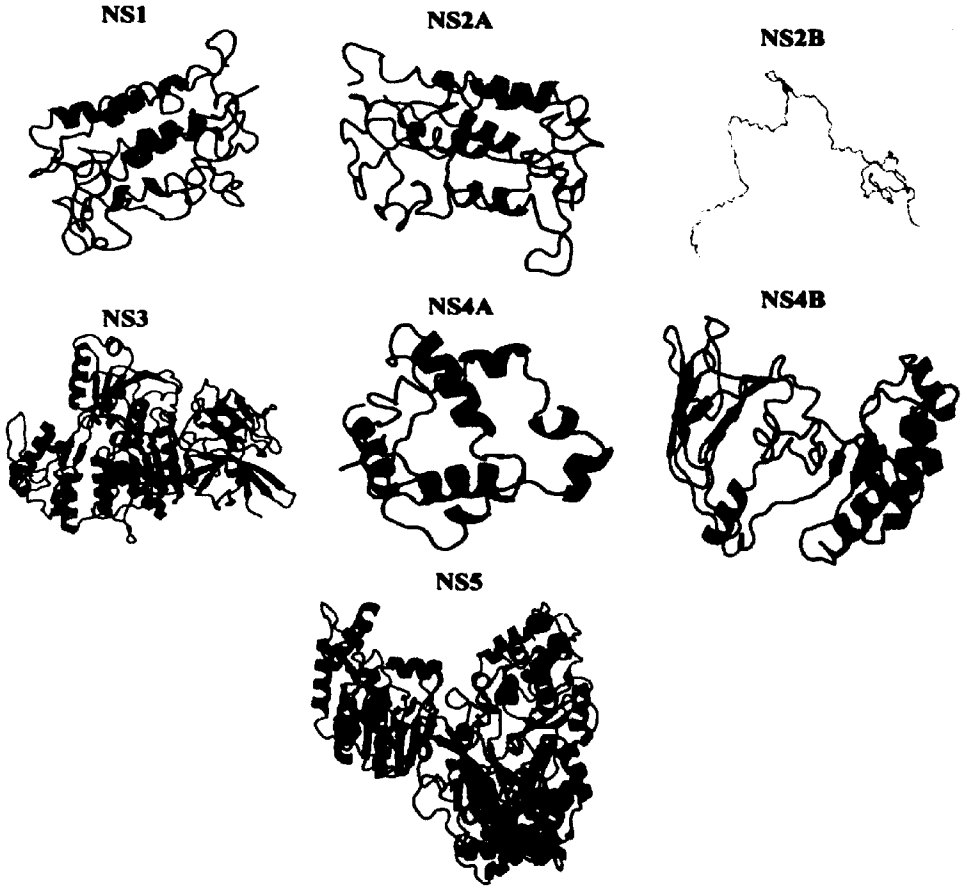


Figure 23: The molecular modelling of the non-structural proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B and NS5) of West Nile virus. The modelling of proteins done in MODELLER 9.10 and molecular visualization was seen in Pymol.

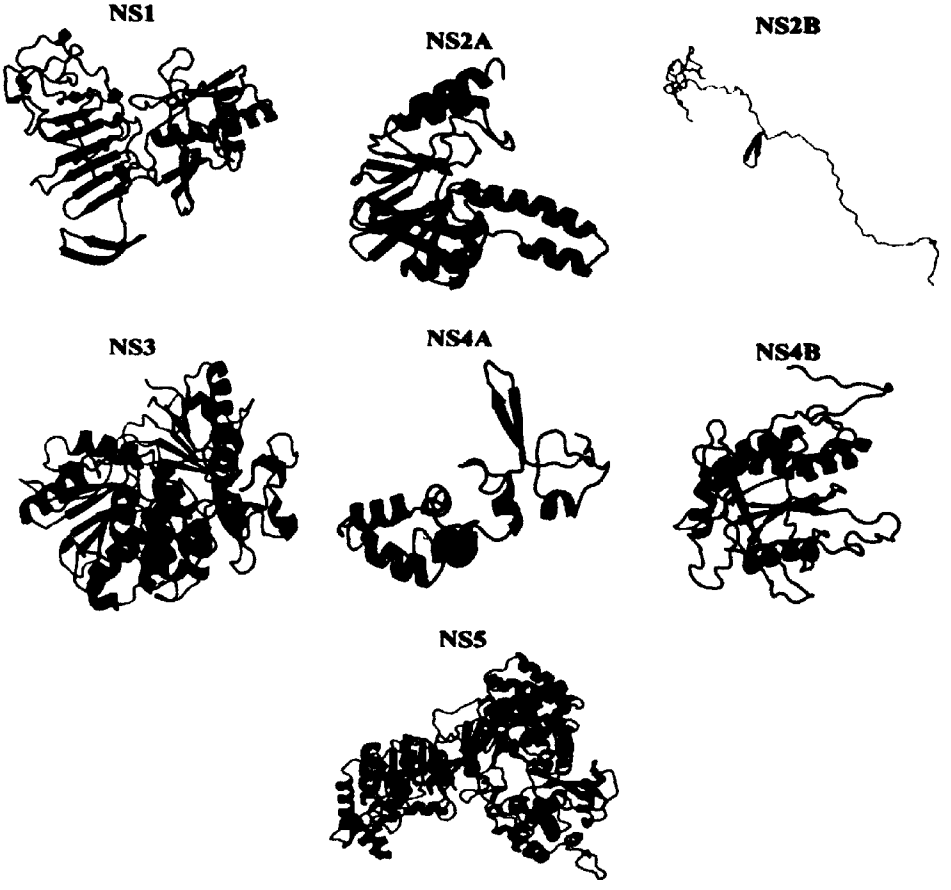


Figure 24: The molecular modelling of the non-structural proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B and NS5) of Yellow Fever Virus. The modelling of proteins done in MODELLER 9.10 and molecular visualization was seen in Pymol.