Journal of Emerging Diseases and Virology

Structural Analysis and Molecular Modeling Studies of Fatty Acids and Peptides Binding with NS2B/NS3 Dengue Protease

NHV Kutumbarao and D Velmurugan*

CAS in Crystallography and Biophysics, University of Madras, Guindy Campus, Chennai, India

*Corresponding author: D Velmurugan, UGC-BSR Faculty, CAS in Crystallography and Biophysics, University of Madras, Guindy Campus, Chennai, India, Tel: 9841075847; **E-mail:** shirai2011@gmail.com

Abstract

Dengue is one of the life threatening diseases. The non-structural protein3 (NS3) is the virus protease essential for the polyprotein processing which requires the presence of ~40 residue hydrophilic domain from NS2B cofactor for its optimal catalytic activity. The complex of NS2B/NS3 active protease is categorized as a trypsin like protease. Design of peptide based drugs towards the protease is one of the widely accepted strategies against this virus. The conformation adopted by NS2B plays an important role in the activity of the protease and also in the binding of the substrate. We have attempted to design peptides against NS2B/NS3 through various approaches. The ligands were selected from different sources, from peptides isolated from edible fishes to natural compounds isolated from papaya leaves. The binding mode of the ligands was studied with respect to different conformations of the protease.

Keywords: DENV; Induced fit docking; Glide; OPLS force field

Abbreviations: DENV: Dengue Virus; WHO: World Health Organization; OPLS: Optimized Potentials for Liquid Simulations.

Introduction

Dengue virus (DENV) belongs to *Flavivirus*. Four antigenetically related serotypes are present, namely DEN-1, DEN-2, DEN-3, DEN-4. The virus is transmitted by *Steomiya aegypti* (Aedes). All the four serotypes are responsible for the invocation of the hemorrhagic fever [1]. The WHO has categorized Dengue as the most important mosquito-borne tropical disease. The symptoms of dengue infections vary from flu-like illness (dengue fever) to dengue shock syndrome and in cases the most severe dengue hemorrhagic fever (severe dengue with bleeding abnormalities). Dengue hemorrhage diseases are life-threatening.

Dengue virus genome is a single-stranded positive-sense RNA [2,3] that consists of 10,723 nucleotides which encodes a single polyprotein precursor which constitutes two categories of proteins, structural and nonstructural proteins. There are three structural proteins (C, prM and E) and a lipid bilayer around the RNA genome [4], where the protein C (core nucleocaspid protein) binds to RNA directly and the protein E (major envelope protein) and protein M (membrane protein) both form the protein outer shell [5]. Seven non-structural proteins (NS) are present, namely, NS1, NS2A, NS2B, NS3, NS4A, NS4B and NS5. The polyprotein precursor undergoes cleavage (co- and post-translationally) to produce mature proteins. The NS3 responsible for this activity, makes the virus active and helps in further replication. The NS3 acts specifically at region for the cleavage activity, the regions are NS2A/NS2B, NS2B/NS3, NS3/NS4A and NS4B/NS5 at the nonstructural protein region [6-10].

The importance of the protease activity in the viral survival and replication and the notion of protease inhibitors being commonly viewed as a potential drug in many cases [11-13] led the research community to design active inhibitors to tackle the dengue infection. The protease

domain NS3 fragment which is 180 amino acid in length is present on the N-terminal of the total 618 residue length multi domain NS3 [14-20]. NS2B (cofactor) is essential for the activation of the NS3 protease activity [16]. The binding of the NS2B might initiate a structural arrangement of the active catalytic triad for the optimal protease activity [21,22]. The protease has specificity for substrate binding, which has been demonstrated by many workers. The results of Niyomrattanakit concludes that the preference for P1 and P2 positions are dibasic residues, with basic or aliphatic residues at P3 and P4 and P1' with smaller or polar residues [23]. The minimal length of the NS3 determines the protease activity. The 47 residue length of the NS2B is determined essential for the protease to be active. The glycine linker connected NS2B to NS3 is soluble and enzymatically active [19,24]. The protease sequences have high sequence similarity within the serotype, the active site residues are conserved over all the serotypes. The sequence alignment of the different protease structures solved from two different serotypes has been represented in the figure 1 [25]. The high similarity between the sequences is evident. This can be observed even in the case when compared with west Nile virus protease also. The DEN2 and DEN3 serotypes proteases can be observed as highly conserved, especially around the active site residues, whose color is in red. The NS2B region too has the similarity between the stereotypes which allows us to model the missing segment of one protease serotype from the other.

The NS2B/NS3 protease is a typical serine protease and the first NS2B/NS3 crystal structure was solved from DEN2 strain at 1.5 Å resolution (PDBID: 2FOM) [26] This is a beta barrel conformation which is similar to chymotrypsin which has active site composed of three major residues, Histidine (HIS), Aspartic acid (ASP) and Serine (SER), which is named as catalytic triad. This structure has a gap in the loop region of the

NS3

```
                              1        10        20        30        40        50        60
                              .        .         .         .         .         .         .
P17763.2:DENV1:   SGVLWDTPSPPEVERAVLDDGIYRILQRGLLGRSQVGVGVFQEGVFHTMWHVTRGAVLMY
3U1I:DENV3:       SGVLWDVPSPPETQKAELEEGVYRIKQGIEGKTQVGVGVQKEGVFHTMWHVTRGAVLTH
2FOM:DENV2:       AGVLWDVPSPPPVGKAELEDGAYRIKQKGILGYSQLGAGVYKEGTFHTMWHVTRGAVLMH
4M9T:DENV2:       AGVLWDVPSPPPVGKAELEDGAYRIKQKGILGYSQLGAGVYKEGTFHTMWHVTRGAVLMH
2M9P:DENV2:       AGVLWDVPSPPPVGKAELEDGAYRIKQKGILGYSQLGAGVYKEGTFHTMWHVTRGAVLMH
Q2YHF2.1:DENV4:   SGALWDVPSPATAQKATLTEGVYRIMQRGLLGRSQVGVGVFHMEGVFHTMWHVTRGSVICH
2FP7:WNV:         ..........L..GDTITGVYRIMTRGLLGSYQAGAGVMVEGVFHTLWHTTRGAALMS
consensus>50:     .gvlwdvpspp.v.kaeledGvYRIkq.GilG.sQvGaG!ykEGvFHT$WHvTrGavlmh

                              70        80        90        100       110       120
                              .         .         .         .         .         .
P17763.2:DENV1:   QGKRLEPGWASVKKDLISYGGGWRFQGSWNAGEEVQVIAVEPGKNPKNVQTAPGTFKTPE
3U1I:DENV3:       NGKRLEPNWASVKKDLISYGGGWRLSAQWQKGEEVQVIAVEPGKNPKNFQTMPGTFQTT
2FOM:DENV2:       KGKRIEPGWADVKKDLISYGGGWKLEGEWKEGEEVQVLALEPGKNPRAVQTKPGLFKTNT
4M9T:DENV2:       KGKRIEPGWADVKKDLISYGGGWKLEGEWKEGEEVQVLALEPGKNPRAVQTKPGLFKTNT
2M9P:DENV2:       KGKRIEPGWADVKKDLISYGGGWKLEGEWKEGEEVQVLALEPGKNPRAVQTKPGLFKTNA
Q2YHF2.1:DENV4:   ETGRLEPGWADVRNDMISYGGGWRLGDKWDKEEDVQVLALEPGKNPKHVQTKPGLFKTLT
2FP7:WNV:         GEGRLDPYWGSVKEDRLGYGGGPWLQHKWNGHDEVQMIVVEPGKNVKNVQTKPGVFKTPE
consensus>50:     .gkRl#PsWadVkkDlisYGGGgWklegeWn.g##VQvlavEPGKNpk.vQTkPGlFkT.t

                              130       140       150       160       170       180
                              .         .         .         .         .         .
P17763.2:DENV1:   GEVGAIALDFKPGTSGSPIVNREGKIVGLYGNGVVTTSGTYVSAIAQAKASQEGPLPEIE
3U1I:DENV3:       GEICAIALDFKPGTSGSPIINREGKVVGLYGNGVVTKNGGYVSGIAQTNAEPDGPTPELE
2FOM:DENV2:       GTICAVSLDFSPGTSGSPIVDKKCKVVGLYGNGVVTRSGAYVSAIANTEKSIED.NPEIE
4M9T:DENV2:       GTIGCVSLDFSPGTSGSPIVDKKCKVVGLYGNGVVTRSGAYVSAIANTEKSIED.NPEIE
2M9P:DENV2:       GTICAVSLDFSPGTSGSPIIDKKCKVVGLYGNGVVTRSGAYVSAIAQTEKSIED.NPEIE
Q2YHF2.1:DENV4:   GEICAVTLDFKPGTSGSPIINKKGKVIGLYGNGVVTKSGGYVSLITQAERIGE.PDYEVD
2FP7:WNV:         GEICAVTLDYPTGTSGSPIVDKNGDVIGLYGNGVIMPNGSYISAIVQGERMEEPAPAGFE
consensus>50:     Ge!Ga!.LD%.pGTSGSPI!#kkGk!!GLYGNGV!t.sG.Y!SaIa#te.s.#..npei#

P17763.2:DENV1:   DEVFRK.
3U1I:DENV3:       EE.....
2FOM:DENV2:       DDIFRK.
4M9T:DENV2:       DDIFRK.
2M9P:DENV2:       .......
Q2YHF2.1:DENV4:   EDIFRK.
2FP7:WNV:         PEMLRKK
consensus>50:     deifrk.
```

NS2B

```
                              1        10        20        30        40
                              .        .         .         .         .
P17763.2:DENV1:   ......ADLSLEKAAEVSWEEEAEHSGASHNILVEVQDDGTMKIKDEERDDTL.
3U1I:DENV3:       ..GPLGSDLTVEKAADVTWEEEAEQTGVSHNLMITVDDDGTMRIKDDETENIL.
2M9P:DENV2:       ...GSAADLELERAADVKWEDQAEISGSSPILSITISEDGSMSIKNEEEEQTL.
4M9T:DENV2:       GSHMLEADLELERAADVRWEEQAEISGSSPILSITISEDGSMSIKNEEEEQTL.
2FOM:DENV2:       GSHMLEADLELERAADVKWEEQAEISGSSPILSITISEDGSMSIKNEEEEQTL.
Q2YHF2.1:DENV4:   ......ADLSLERAANVQWDEMADITGSSPIIEVKQDEDGSFSIRDVEETNMI.
2FP7:WNV:         GSHMLETDMWIERTADITWESDAEITGSSERVDVRLDDDGNFQLMNDPGAPWKG
consensus>50:     ....l.aD$.lEraA#!.W#eqA#isGsSpil.!tid#DGsmsik#eeeeqtl.
```
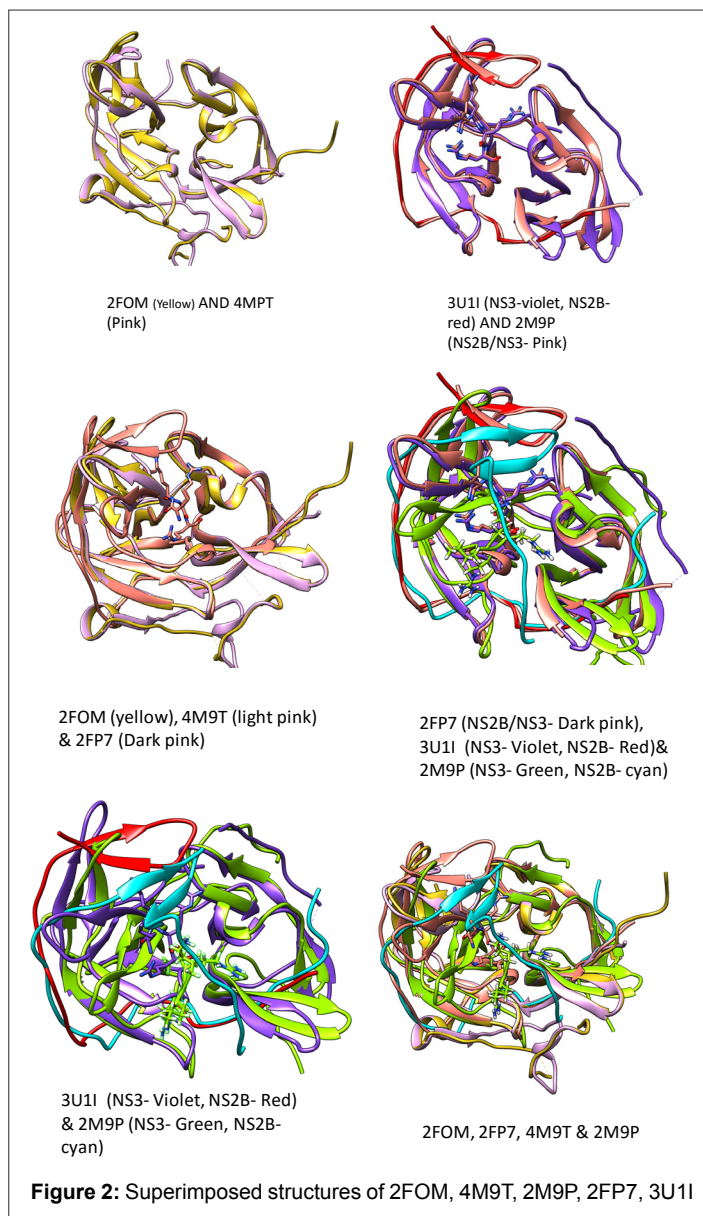
Figure 1: The sequence comparison of the protease NS3 and NS2B

NS2B. Many of the subsequent structures solved have missing loops, one structure solved from DEN2 (PDBID: 4M9T) [27] with reported allosteric site has the trace of the loop and the orientation of the NS2B was similar to that of 2FOM structure. But a solution structure (PDBID:2M9P) recently deposited in PDB has an inhibitor bound in the active site and the NS2B region shows a major conformational change which was similar to the conformation of protease (PDBID:3U1I) solved from DEN3 [28] .The 2FOM structure which was the earliest solved structure has been recently reported to be in the inactive conformation. The orientation of the NS2B fragment in this conformation is compared to that of the structure 4M9T which was a mutant structure for A125C, and this structure is reported to have helped in the identification of the allosteric region (ALA 125) in the protease. The movement of the loop is identified to have an influential role. The conformations of the loop120 (117-122) and loop150 (153–164) [27] are crucial and influence the orientation of the NS2B fragment. The different positions of the loop with respect to the NS2B orientation and also the place of the ligand can be seen in figure 2. The movement of these loops is indeed linked to the binding mode of the ligand. This can be seen from the binding of different ligands shown in figure 2. The RMSD (Root Mean Square deviation) and orientations of superimposed structure 2FOM with 3U1I is 0.6 Å (127 atom pairs included) where as there is a deviation of 1.03 Å with 2M9P (only 80 atom pairs included). The RMSD between 3U1I and 2M9P is 1.18 Å (88 atom pairs included). There is a conformational similarity between 3U1I (DEN3) and the protease from West Nile virus, 2FP7 [26]. The positioning of the loop 120 is further deviated in the recent solution structure and the NS2B also. This can be inferred as the effect brought by different binding modes of the ligands. The various superimpositions of the structures are presented in figure 2. As the conformation adopted by the protease in 2FOM structure is inactive [29], many workers have used the subsequent structure such as 3U1I from DEN3 for modeling studies. Molecular docking and simulation studies were undertaken to understand the different binding modes of the ligands and the effect of placement of the loops for the ligand binding. The previous modeling studies were carried out using 2FOM as the target and we have analyzed the recent structures deposited and carried out the modeling analysis which lead to the binding mode of the few peptides similar to the binding mode of ligand seen in the 2M9P solution structure. The structure used for modeling from NMR (Nuclear Magnetic Resonance spectroscopy) studies is the one with the least energy.

Materials and Methods

The proteins were downloaded from Protein Data Bank (PDB). The peptides were modeled using PyMOL. The three dimensional structures of the natural products from papaya leaves were downloaded from

Figure 2: Superimposed structures of 2FOM, 4M9T, 2M9P, 2FP7, 3U1I

ligand and the protein as flexible. Induced fit (Glide XP) module which possesses flexible docking option was used for docking of ligands with protein. The grid was specified for the site of docking by specifying the active site residues, in this case, the catalytic triad residues. Grid of 20 Å along each edge is specified for the calculations to be performed. The resulting output file was analyzed and the best pose was considered for molecular simulation studies. The molecular dynamic simulations were carried out using AMBER 12 (Assisted Model Building with Energy Refinement) [31]. The molecular dynamic simulation helps to analyze the interactions and behavior of the ligand in dynamic state over a time scale. This also helps us to know the stability of the protein-ligand complex. AMBER FF99SB force field was used for the parameterization of the protein molecule. TIP3P water box was used for the solvation of the complex and charge neutralization was carried out using Na^+ and Cl^- ions. The total complex with water molecules was minimized first and then equilibration was carried out until the system reaches a stable temperature and pressure.

Results and Discussion

Compounds from papaya leaf and synthetic compounds

Extracts from different plant sources are known to have medicinal importance. There are many successful cases where compounds and secondary metabolites from the plants have been proved to have anti-microbial activity. In the recent outbreak of dengue in India, there were reports that many medical practitioners have used papaya leaf extract to cure dengue fever and cases of successful treatment were also seen. Extracts from the papaya leaf have been reported to have an anti-dengue activity [32]. In the light of this we have analyzed the papaya leaf extract using GCMS (Gas Chromatography Mass Spectrometry) and found three secondary metabolites, oleic, stearic and palmatic acids as major constituents. We have then carried out molecular modeling studies of the three compounds towards the NS2B/NS3 protease, the interactions and the glide score were relatively higher and the interaction with active site residues were favorable when the docking was carried out with 2M9P as target compared with 2FOM as the target. The docking score, glide energy and interaction diagrams are presented. All the three compounds possess interactions with the catalytic residues, with one H-bond interaction and some non-bonded interactions also. The region of the binding at the catalytic site for oleic and palmatic acid is similar but it is different in the case of the stearic acid. The information regarding the score, energy and interaction are shown in table1 and the binding sites are shown in the figures 3a and 3b. The ligands oleic acid and palmatic acid were bound in a similar binding mode as in the co-crystal structure, where as the stearic acid has a different binding. The binding orientation of the two ligands which were similar to the co-crystal is represented in figure 3c. The ligand is shown in stick and the protein residues are shown in line format, with the active site residues highlighted in cyan colour. The compounds showed an improved score and energy in the binding with the protein in the modeled NMR structure, where the full structure is modeled with NS2B fragments.

Public chemical database (Pubchem). The three dimensional structures of synthetic compounds were obtained from crystallographic studies. The structure of the ligands has to be optimized in order to rectify steric clashes and also to calculate the minimum potential energy. So, prior to docking, all the ligands were minimized using OPLS 2005 force-field (Optimized Potentials for Liquid Simulations) in the Impact module of Schrödinger 2009. During minimization, the ligands were subjected first to steepest descent for 1000 iterative cycles available in the Impact minimization module. This is a primary step, where the initial ligand geometry with steric clashes effect can be treated properly. This was then followed by conjugate gradient for 5000 cycles which makes good convergence in the structure with respect to energy and gradient. The output of this was chosen for the docking [30]. The protein was minimized using protein preparation wizard where addition of H atoms and bond order were adjusted and further energy minimization was carried out using OPLS2005 force-field. Molecular docking helps in identifying energetically and geometrically favorable binding pose of a ligand bound to the protein. Out of different types of docking, Induced fit docking possess more advantage. This method of docking helps to treat both the

Compound	Docking Score	Glide Energy Kcal/mole	Interaction D…A	Distance (Å)
Oleic acid	-4.874	-45.371	UNK (O)…ILE97(O) HIS112(N)…UNK(O)	2.63 3.06
Stearic acid	-4.576	-38.793	UNK (O)…GLU(O)230	2.73
Palmatic acid	-4.6	-40.543	UNK (O)…SER(O)196	2.78

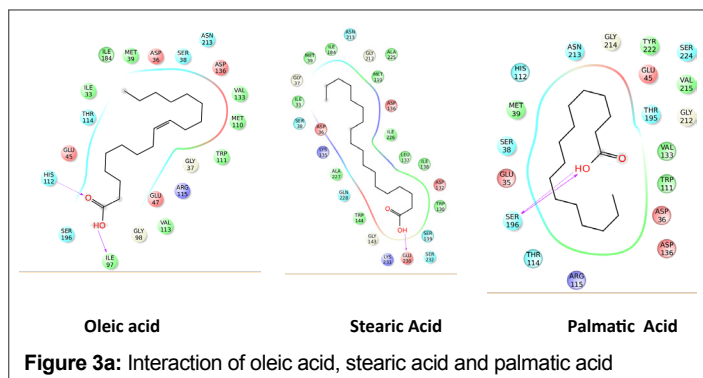**Table 1:** Induced fit docking results of compounds from papaya leaf

Figure 3a: Interaction of oleic acid, stearic acid and palmatic acid
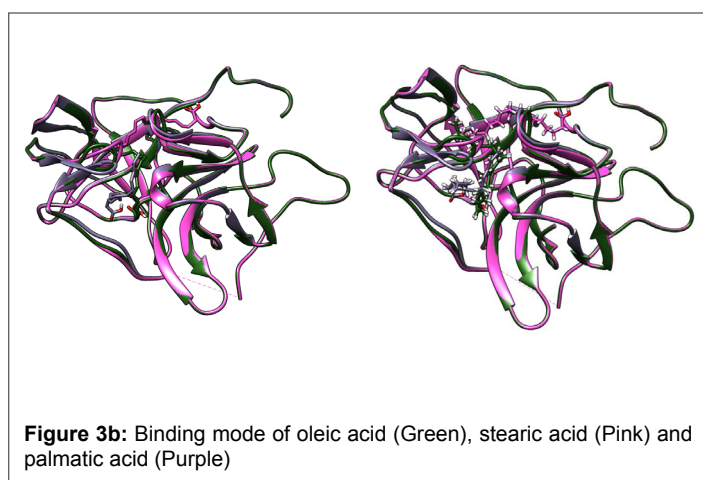

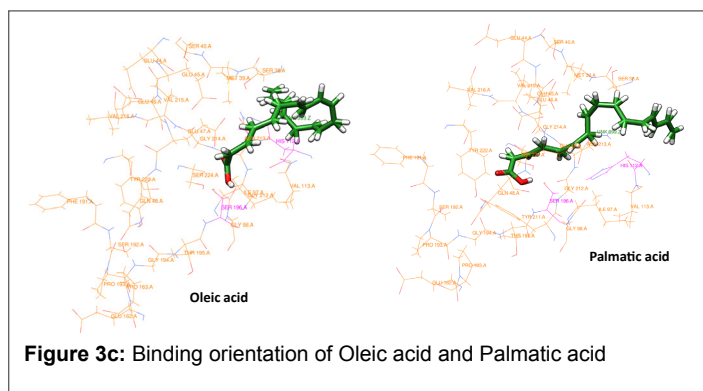**Figure 3b:** Binding mode of oleic acid (Green), stearic acid (Pink) and palmatic acid (Purple)

Figure 3c: Binding orientation of Oleic acid and Palmatic acid

Figure 4: Binding mode of different peptides superimposed with 2M9P (violet)

Table 2: Induced fit docking results of Peptides

Peptide	Docking Score	Glide Energy Kcal/mole
SMHG	-10.5	-79.9
GHMS	-11.5	-73.4
ASIF	-11.5	-80.1
CREAC	-13.1	-89.3
EGPR	-9.01	-80.1
KGPE	-8.6	-81.8
LEVKP	-13.6	-83.8
LFCA	-9.2	-79.1
RPE	-10.3	-65.1

The compounds which were synthesized by our collaborators and reported [33] were also subjected to modeling using 3U1I as template. These compounds also showed improved binding, and they were also subjected to dynamic simulations. The analysis showed that the compounds were bound and showed a better interaction during the course of the trajectory.

Peptides as inhibitors

The use of peptides as favorable drugs than the synthetic compounds has gained much interest, as they are less toxic with minimum side effects. There have been many efforts to design peptides based on the active site structure or based on the substrate. Many peptide leads have been reported in the recent times as potent inhibitors against the dengue protease [34]. These peptides are also end modified to incorporate the influence of the functional groups. From our previous modeling studies we have reported peptides SHMG, GHMS isolated from edible fishes, and designed peptides
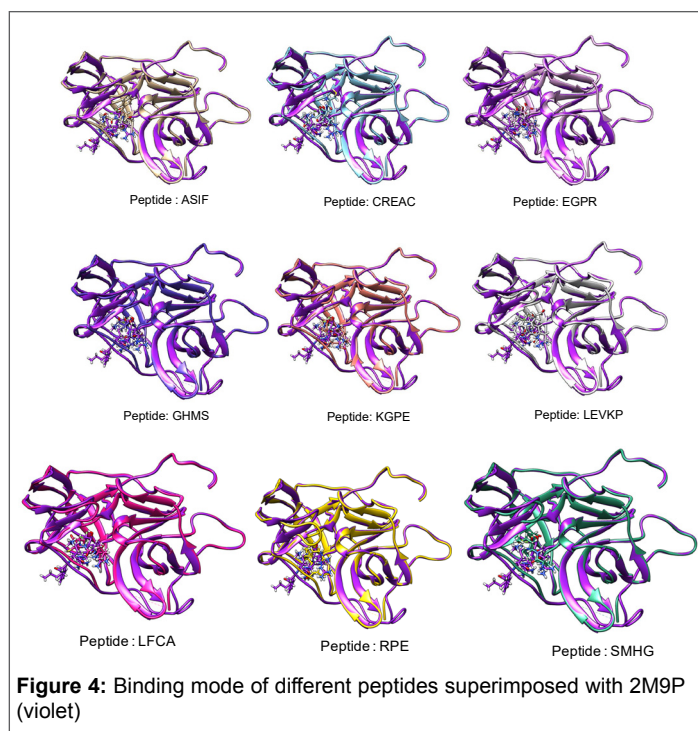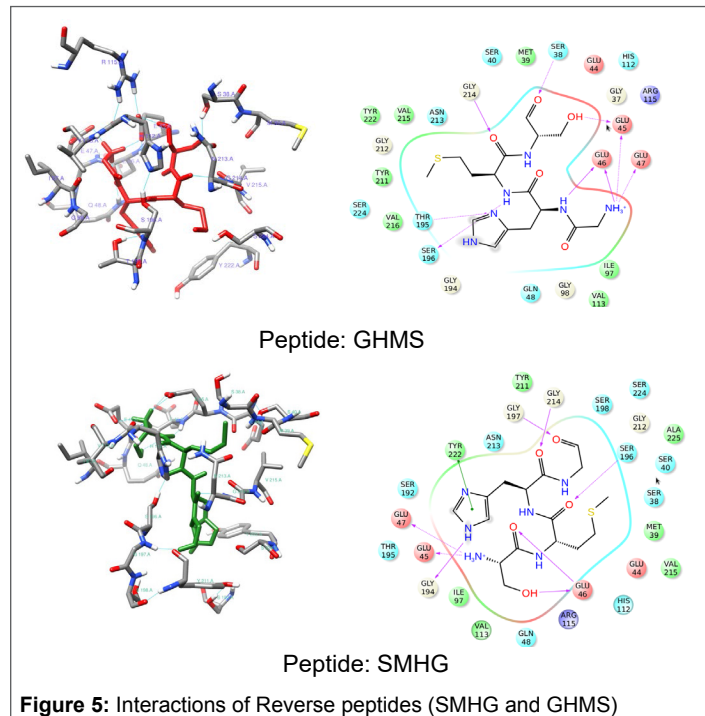
which can bind with good energy and score [35]. Peptides which showed good results have now been subjected to docking with the 2M9P as target to see if the interaction has been subjected to any change. The binding of the peptides to the target was found to be better with improved score and energy values. The docking score and energy values are tabulated (Table 2) with the interactions shown in figure 4. The binding site of the majority of the peptides is similar to that of the co-crystal ligand in the solution structure. With an effect on the energy but with increase in docking score, the binding of other peptides is not only at the active site of NS3 but their interactions can also be seen with residues of NS2B. These possess hydrogen mediated interactions with SER and HIS and maintain non-bonded interactions with the other active site residues. The presence of proline and glycine in the peptide as reported in the previous modeling papers show an affinity with additional interactions. The orientation of the reverse peptides makes these to interact with both the chains. In view of the peculiar nature of its sequence and its availability in nature (these peptides are isolated from edible fishes), the reverse peptides were subjected to simulation studies. The target bound peptides were subjected to molecular dynamic simulation for 30 ns and showed a consistent binding throughout the trajectory time (data not shown). The position of the peptides and their superimposed information of the individual peptides with the 2M9P are given in the figures 5,6. It can be observed that
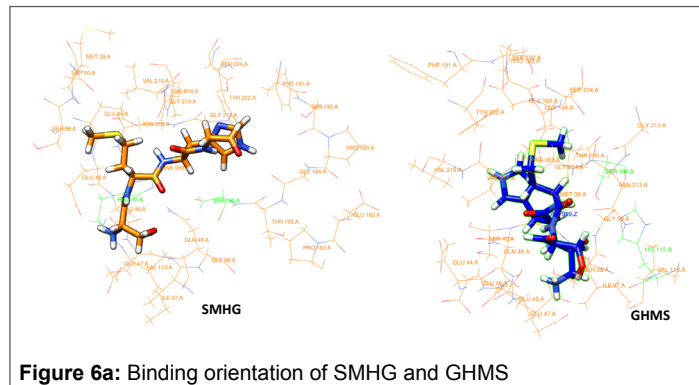
Peptide	Interaction		Distance (Å)
SMHG	GLU 46.A N	UNK 999.Z O	2.8
	SER 196.A OG	UNK 999.Z O	2.8
	GLY 197.A N	UNK 999.Z O	3
	GLY 214.A N	UNK 999.Z O	2.7
	UNK 999.Z N	GLU 45.A OE1	3.5
	UNK 999.Z N	GLU 45.A OE2	2.9
	UNK 999.Z N	GLU 47.A OE2	2.6
	UNK 999.Z NE2	GLY 194.A O	3.1
	UNK 999.Z OG	GLU 46.A O	3
GHMS	SER 38.A OG	UNK 999.Z O	2.7
	THR 195.A OG1	UNK 999.Z ND1	3
	GLY 214.A N	UNK 999.Z O	2.8
	UNK 999.Z N	GLU 45.A OE2	2.6
	UNK 999.Z N	GLU 46.A O	2.6
	UNK 999.Z N	GLU 46.A O	2.9
	UNK 999.Z N	GLU 47.A OE2	2.8
	UNK 999.Z N	SER 196.A OG	2.7
	UNK 999.Z OG	GLU 45.A OE1	3
ASIF	SER 38.A OG	UNK 999.Z OG	2.8
	GLU 46.A N	UNK 999.Z O	2.8
	ARG 115.A NE	UNK 999.Z O	2.7
	UNK 999.Z N	GLU 44.A O	2.9
	UNK 999.Z N	GLU 44.A OE1	2.6
	UNK 999.Z N	GLU 45.A OE2	2.8
	UNK 999.Z N	ASN 213.A OD1	3
	UNK 999.Z OG	GLU 45.A OE1	3.2
CREAC	SER 38.A OG	UNK 999.Z O	2.7
	ASN 213.A ND2	UNK 999.Z O	2.7
	UNK 999.Z N	ASP 36.A OD1	2.6
	UNK 999.Z N	GLY 37.A O	3.1
	UNK 999.Z N	GLY 214.A O	2.9
	UNK 999.Z NE	HIS 112.A O	3.3
	UNK 999.Z NE	VAL 113.A O	3.4
	UNK 999.Z NH1	GLU 45.A OE2	2.7
	UNK 999.Z NH1	GLU 47.A OE2	2.7
	UNK 999.Z NH2	GLU 47.A OE2	2.8
	UNK 999.Z NH2	VAL 113.A O	2.9
	UNK 999.Z SG	GLY 214.A O	3.5
EGPR	HIS 112.A ND1	UNK 999.Z O	2.7
	ARG 115.A NE	UNK 999.Z OE2	2.8
	ARG 115.A NH2	UNK 999.Z OE1	2.7
	SER 196.A N	UNK 999.Z O	2.8
	GLY 214.A N	UNK 999.Z O	2.9
	UNK 999.Z N	GLU 45.A OE1	2.8
	UNK 999.Z N	GLU 45.A OE2	3.3
	UNK 999.Z N	GLU 46.A O	2.8
	UNK 999.Z NE	GLU 47.A OE2	3.1
	UNK 999.Z NH1	GLU 47.A OE2	2.6
	UNK 999.Z NH2	HIS 112.A O	2.7
KGPE	GLN 48.A NE2	UNK 999.Z O	2.8
	GLY 98.A N	UNK 999.Z OE2	2.7
	SER 196.A N	UNK 999.Z O	2.9
	GLY 214.A N	UNK 999.Z O	3.2
	UNK 999.Z N	GLU 45.A OE1	3
	UNK 999.Z N	GLU 45.A OE2	2.8
	UNK 999.Z N	GLU 45.A OE2	2.9
	UNK 999.Z N	HIS 112.A O	2.9
	UNK 999.Z NZ	GLU 45.A OE1	2.9
	UNK 999.Z NZ	GLU 47.A OE1	3.5
	UNK 999.Z NZ	GLU 47.A OE2	2.5

LEVKP	GLU 46.A N	UNK 999.Z O	2.9
	GLN 48.A NE2	UNK 999.Z OE2	2.8
	ARG 115.A NH1	UNK 999.Z O	2.9
	SER 196.A N	UNK 999.Z OE1	3
	ASN 213.A ND2	UNK 999.Z O	3.4
	UNK 999.Z N	SER 196.A OG	3.1
	UNK 999.Z NZ	GLU 46.A O	2.9
	UNK 999.Z NZ	GLU 47.A OE1	2.9
LFCA	HIS 112.A ND1	UNK 999.Z O	2.7
	SER 224.A OG	UNK 999.Z O	2.9
	UNK 999.Z N	GLU 45.A OE2	2.8
	UNK 999.Z N	GLU 47.A OE2	2.9
	UNK 999.Z N	ASN 213.A OD1	3.1
	UNK 999.Z S	GGLU 44.A O	3.8
RPE	GLN 48.A NE2	UNK 999.Z O	2.8
	GLN 96.A NE2	UNK 999.Z OE1	2.7
	THR 195.A OG1	UNK 999.Z OE1	3.4
	THR 195.A OG1	UNK 999.Z OE2	2.7
	SER 196.A N	UNK 999.Z O	3
	UNK 999.Z N	GLU 44.A OE1	2.8
	UNK 999.Z N	UNK 999.Z O	2.7
	UNK 999.Z NE	GLY 37.A O	3.3
	UNK 999.Z NH1	GLU 47.A OE1	2.8
	UNK 999.Z NH1	HIS 112.A O	2.5
	UNK 999.Z NH2	GLU 47.A OE2	2.7

Table 3: Interaction details of the peptides with protease

Peptide: GHMS

Peptide: SMHG

Figure 5: Interactions of Reverse peptides (SMHG and GHMS)

Figure 6a: Binding orientation of SMHG and GHMS

are represented in the cyan colour, the ligands are represented in stick model with the protein residues been projected as line with three letter indicator with residue name and number (table 3).

Conclusion

Our attempt to design compounds against dengue virus protease has not been a straight forward task. Even though the main architecture of the dengue protease is similar to serine protease, the presence of the active site on the surface of the protein, hydrophobic pocket and the influence of the NS2B cofactor have a large influence on the ligand binding. We have designed a series of peptides specific to the protease, through substrate based and active site based approach. Their binding modes towards the protease were analyzed through modeling studies and are validated by subjecting the complex to molecular dynamic simulations. This approach may help to find out not only a static low energetic structure but also a stable complex. This can help in completing the *in silico* method for the identification of the best possible ligand as
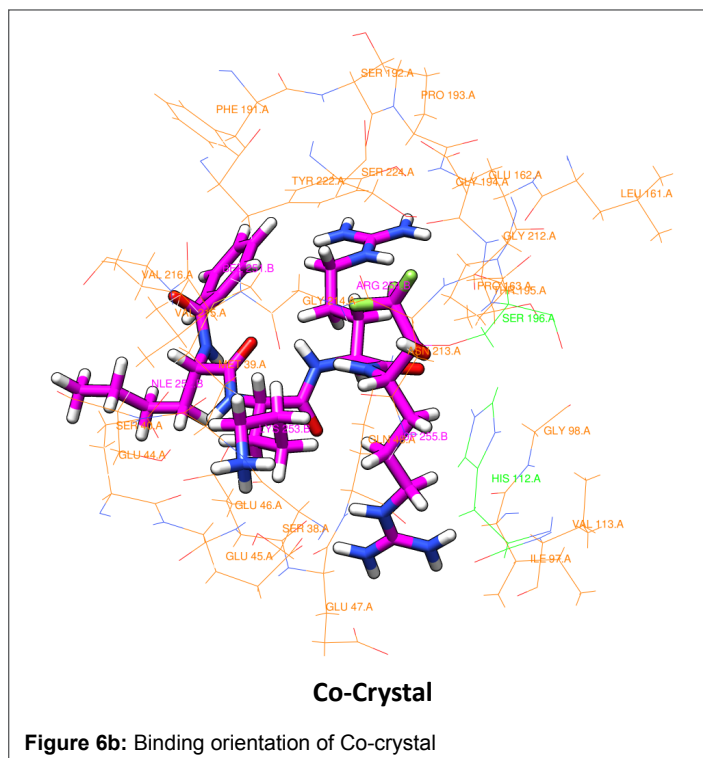
the binding pocket of the peptides, GHMS and SMHG are similar to that of the co-crystal. The binding of the ligand with the Histidine and Serine amino acids can be found in both the peptides as seen in the co-crystal. The GHMS has a similar orientation in the pocket as in the case of co-crystal. Figures 6a and 6b show the binding orientations of the peptides and co-crystal ligand in the pocket respectively. The active site residues

Co-Crystal

Figure 6b: Binding orientation of Co-crystal

an inhibitor. From this particular study, one can observe that the binding efficiency of ligands is influenced by the protein state and the crucial residues, which at times won't show up in the crystallographic studies. The modeling of the missing residues with help of conserved and similar counterpart will help in choosing the better ligand. Docking and dynamic studies carried out with the structure reported from NMR studies also confirmed improvement in the binding affinity and also the mode of binding. The molecular dynamics further helped us in finding the dynamic nature of the ligand also and its energetics in binding with the active site residues. Encouraged by the above results attempts are underway for co-crystallization of NS2B/NS3 with peptides and also with Oleic acid and Palmatic acid.

Acknowledgement

References

1. Murray N E A, Quam M B, Wilder-Smith A (2013) Epidemiology of dengue: past, present and future prospects. Clin Epidemiol 5: 299-309.

2. Gubler D J, Kuno G, Markoff L (2006) Flaviviruses. In: Knipe DM, Howley PM (Eds) Fields Virology, Chapter 34, 5th edition. Lippincott Williams & Wilkins, USA, 1153-1252.

3. Kuno G, Chang GJ, Tsuchiya KR, Karabatsos N, Cropp CB (1998) Phylogeny of the genus Flavivirus. J Virol 72: 73-83.

4. Lindenbach BD, Thiel H-J, Rice CM (2007) Flaviviridae: the viruses and their replication. In: Knipe DM, Howley PM (eds), Fields virology, 5th edition. Lippincott Williams and Wilkins, Philadelphia, USA, 1102-1152.

5. Kuhn R J, Zhang W, Rossmann M G, Pletnev S V, Corver J, et al. (2002) Structure of dengue virus: implications for flavivirus organization, maturation, and fusion. Cell 108: 717-725.

6. Chambers TJ, Weir RC, Grakoui A, McCourt DW, Bazan JF, (1990) Evidence that the N-terminal domain of non-structural protein NS3 from yellow fever virus is a serine protease responsible for site-specific cleavages in the viral polyprotein. Proc Natl Acad Sci USA 87: 8898-8902.

7. Lin C, Amberg SM, Chambers TJ, Rice CM (1993) Cleavage at a novel site in the NS4A region by the yellow fever virus NS2B-3 proteinase is a prerequisite for processing at the downstream 4A/4B signalase site. J Virol 67: 2327-2335.

8. Lobigs M (1993) Flavivirus premembrane protein cleavage and spike heterodimer secretion require the function of the viral proteinase NS3. Proc Natl Acad Sci USA 90: 6218-6222.

9. Preugschat F, Yao CW, Strauss JH (1990) *In vitro* processing of dengue virus type 2 non-structural proteins NS2A, NS2B, and NS3. J Virol 64: 4364-4374.

10. Teo K F, Wright P J (1997) Internal proteolysis of the NS3 protein specified by dengue virus 2. J Gen Virol 78: 337-341.

11. Tomlinson SM, Malmstrom RD, Watowich SJ (2009) New approaches to structure-based discovery of dengue protease inhibitors. Infect Disord Drug Targets 9: 327-343.

12. Hsu JT, Wang HC, Chen GW, Shih SR (2006) Antiviral drug discovery targeting to viral proteases. Curr Pharm Des 12: 1301-1314.

13. Wlodawer A, Vondrasek J (1998) Inhibitors of Hiv-1 Protease: A Major Success of Structure-Assisted Drug Design. Annu Rev Biophys Biomol Struct 27: 249-284.

14. Lescar J, Luo D, Xu T, Sampath A, Lim S P, et al. (2008) Towards the design of antiviral inhibitors against flaviviruses: the case for the multifunctional NS3 protein from Dengue virus as a target. Antiviral Res 80: 94-101.

15. Arias CF, Preugschat F, Strauss JH (1993) Dengue 2 virus NS2B and NS3 form a stable complex that can cleave NS3 within the helicase domain. Virology 193: 888-899.

16. Falgout B, Miller RH, Lai CJ (1993) Deletion analysis of dengue virus type 4 nonstructural protein NS2B: identification of a domain required for NS2B-NS3 protease activity. J Virol 67: 2034-2042.

17. Li H, Clum S, You S, Ebner KS, Padmanabhan R (1999) The serine protease and RNA-stimulated nucleoside triphosphatase and RNA helicase functional domains of dengue virus type 2 NS3 protein converge within a region of 20 amino-acids. J Virol 73: 3108-3116.

18. Yusof R, Clum S, Wetzel M, Krishna Murthy HM, Padmanabhan R (2000) Purified NS2B/NS3 serine protease of dengue virus type 2 exhibits cofactor NS2B dependence for cleavage of substrates with dibasic amino acids in vitro. J Biol Chem 275: 9963-9969.

19. Leung D, Schroder K, White H, Fang NX, Stoermer MJ, et al. (2001) Activity of recombinant dengue 2 virus NS3 protease in the presence of a truncated NS2B co-factor, small peptide substrates, and inhibitors. J Biol Chem 276: 45762-45771.

20. Li J, Lim SP, Beer D, Patel V, Wen DY, et al. (2005) Functional profiling of recombinant NS3 proteases from all four serotypes of dengue virus using tetrapeptide and octapeptide substrate libraries. J Biol Chem 280: 28766-28774.

21. Niyomrattanakit P, Winoyanuwattikun P, Chanprapaph S, Angsuthanasombat C, Panyim S, et al. (2004) Identification of residues in the dengue virus type 2 NS2B cofactor that are critical for NS3 protease activation. J Virol 78: 13708-13716.

22. Barbato G, Cicero DO, Nardi MC, Steinkuhler C, Cortese R, et al. (1999) The solution structure of the N-terminal proteinase domain of the hepatitis C virus (HCV) NS3 protein provides new insights into its activation and catalytic mechanism. J Mol Biol 289: 371-384.

23. Niyomrattanakit P, Yahorava S, Mutule I, Mutulis F, Petrovska R, et al. (2006) Probing the substrate specificity of the dengue virus type 2 NS3 serine protease by using internally quenched fluorescent peptides. Biochem J 397: 203-211.

24. Nall TA, Chappell KJ, Stoermer MJ, Fang NX, Tyndall JD, et al. (2004) Enzymatic characterization and homology model of a catalytically active recombinant West Nile virus NS3 protease. J Biol Chem 279: 48535-48542.

25. Corpet F (1988) Multiple sequence alignment with hierarchical clustering. Nucl Acids Res 16: 10881-10890.

26. Erbel P, Schiering N, D'Arcy A, Renatus M, Kroemer M, et al. (2006) Structural basis for the activation of flaviviral NS3 proteases from dengue and West Nile virus. Nature Struct Molecular Biol 13: 372-373.

27. Yildiz M, Ghosh S, Bell JA, Sherman W, Hardy JA (2013) Allosteric Inhibition of the NS2B-NS3 Protease from Dengue Virus. ACS Chem Biol 8: 2744-2752.

28. Noble CG, Seh CC, Chao AT, Shi PY (2012) Ligand-bound structures of the dengue virus protease reveal the active conformation. J Virol 86: 438-446.

29. de la Cruz L, Nguyen TH, Ozawa K, Shin J, Graham B, et al. (2011) Binding of low molecular weight inhibitors promotes large conformational changes in the dengue virus NS2B-NS3 protease: fold analysis by pseudocontact shifts. J Am Chem Soc 133: 19205-19215.

30. The Protein Preparation Wizard (2009) Protein Preparation Guide, Chapter 2, Schrödinger, LLC.

31. Case DA, Cheatham TE, Darden T, Gohlke H, Luo R, et al. (2005) The Amber biomolecular simulation programs. J Comput Chem 26: 1668-1688.

32. Subenthiran S, Choon TC, Cheong KC, Thayan R, Teck M K, et al. (2013) Carica papaya Leaves Juice Significantly Accelerates the Rate of Increase in Platelet Count among Patients with Dengue Fever and Dengue Haemorrhagic Fever. Evidence-Based Complementaryn Altern Med.

33. Timiri AK, Selvarasu S, Kesherwani M, Vijayan V, Sinha BN, et al. (2015) Synthesis and molecular modelling studies of novel sulphonamide derivatives as dengue virus 2 protease inhibitors. Bioorg Chem 62: 74-82.

34. Luo D, Vasudevan SG, Lescar J (2015) The flavivirus NS2B-NS3 protease-helicase as a target for antiviral drug development. Antiviral Res 118: 148-158.

35. Velmurugan D, Mythily U, Kutumbarao (2014) Design and Docking Studies of Peptide Inhibitors as Potential Antiviral Drugs for Dengue Virus Ns2b/Ns3 Protease. Protein Pept Lett 21: 815-827.