Bangladesh J Microbiol, Volume 23, Number 2, December 2006, pp 107-113

Original Article

Computational Analyses of NS3 Serine Protease of Dengue Virus

Jesmin* and Nazlee Sharmin

Department of Genetic Engineering & Biotechnology, University of Dhaka, Dhaka 1000, Bangladesh

[Received 07 October 2006; Accepted 12 November 2006]

Dengue virus (DENV), a mosquito-borne *Flavivirus,* is an emerging global health threat. A number of studies have already revealed that the non-structural NS3 serine protease is required for the maturation of the viral polyprotein and thus is a promising target for the development of antiviral inhibitors. However, the residues and other structural elements that play a role in the enzyme-mediated maturation process of DENV by NS3 have yet to be definitely assigned. Identification of the binding site and the actual environment around the active site pocket are still open to questions. To elucidate the functions of DENV NS3 and in particular, for a better understanding of the active site pocket of the enzyme, a 3D model of DEN2 NS3 serine protease to locate its key catalytic residues has been proposed. From computational comparative analyses of sequences and structures of related NS3 serine proteases of the *Flaviviridae* family, the charge distribution and electrostatic potential in and around the active site pocket of DEN2-NS3 serine protease have also been predicted. This proposed model would facilitate future studies for better rationalize the environment of the substrate-binding pocket and thus stimulate more rationally designed structure-function studies aimed at elucidating the role of this enzyme in virus maturation.

Keywords: Dengue virus (DENV), *Flaviviridae*, NS3 serine protease, Polyprotein

Introduction

Dengue is an infectious viral disease endemic to tropical areas worldwide that put an estimated 2.5 billion people at risk[1]. Dengue virus (DENV), a member of the *Flaviviridae* family, is an arthropod-borne human pathogen transmitted by female *Aede aegypti* mosquitoes, which causes a spectrum of clinical symptoms ranging from mild, uncomplicated fever to severe forms of the disease, dengue hemorrhagic fever and dengue shock syndrome[2]. Currently four serotypes (DEN1-4) of Dengue virus exist. Research has shown that infection with one serotype of the virus makes the patient up to 15 times more likely to develop dengue hemorrhagic fever (DHF) if infected for a second time with a different serotype. Infection with one serotype does not provide immunity against the remaining serotypes. Because of this complication, a dengue vaccine against one serotype has been shown to make patients more susceptible to contracting DHF upon exposure to one of the remaining three serotypes[3]. At present, there are neither vaccines nor other treatments available to prevent or cure this disease; the only means of controlling the spread of DENV is an attempt to control the mosquito vector population. These facts emphasize the need for a better understanding of the mechanism of viral infection to combat this disease.

DENV has been the subject of numerous studies. All the four serotypes of DENV have a single-stranded positive sense-stranded RNA (+RNA) genome. The only protein product it translates is a full-length precursor polyprotein. However, post-translational proteolytic cleavage of the precursor protein results in the creation of three structural proteins - capsid (C), membrane (M), envelope (E) and seven non-structural proteins - NS1, NS2a, NS2b, NS3, NS4a, NS4b and NS5[3]. Among the four serotypes, dengue virus type 2 (DEN2) is the most prevalent one. The viral polyprotein and the proteolytic cleavage sites are shown in Figure 1. Two proteases play key roles in proteolytic cleavage process - one of host origin and the other of the viral. Processing at the NS1-NS2A junction is mediated by a host protease within the endoplasmic reticulum that may be either signal peptidase or another unknown enzyme[4]. The remaining cleavages in the polyprotein precursor are mediated by the viral chymotrypsin-like serine protease, encoded within the amino-terminal 180 amino acid residues of NS3 protease[4]. NS3 is essential in host-infected cells for post-translational processing of a viral polypeptide precursor into the structural and functional viral proteins. Flaviviral NS3 protein includes the domains that are necessary to carry out protease, nucleoside triphosphatase (NTPase) and helicase activities[5-8]. This proteases domain identified, based on the sequence similarity of NS3 with several viral and cellular serine proteases[9-10], is a potent target for the development of therapeutics against the virus[3].

Corresponding author:
Dr. Jesmin, Assistant Professor, Department of Genetic Engineering & Biotechnology, University of Dhaka, Dhaka 1000, Bangladesh
Tel (Office): (02) 9661920/7819; Tel (Home): (02) 9662222; E-mail: mailjesmin@yahoo.com

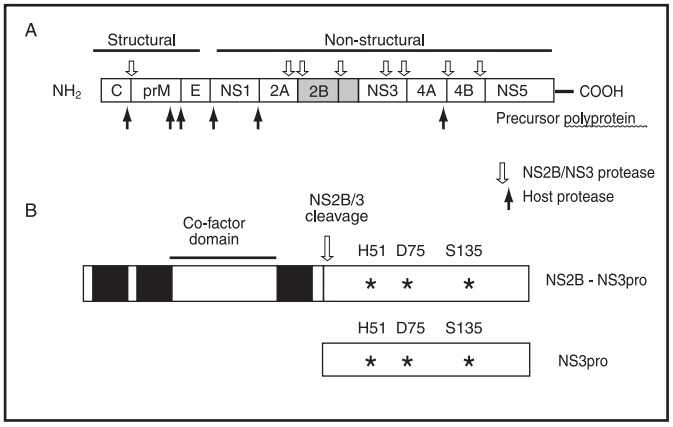


Figure 1. *Polyprotein precursor and the NS2B. NS3pro protease structure of Flavivirus. A. Proteolytic cleavage sites in the viral polyprotein by host cell proteases and the virus-encoded NS2B/NS3 protease. B. Structure of the NS3 protease as generated by splicing (adapted from Chanprapaph et al.[3]).*

Recently, it has also been established that active NS3 serine protease is an essential requirement for maturation of the active virus in DEN3[11]. Another study also confers the involvement of a NS3 serine protease in DEN2[3]. Crystal structures of the related NS3pro from hepatitis C virus (HCV) have been used successfully for drug discovery[12], and a structure of DEN NS3pro lacking the NS2B (1BEF) cofactor is available[13]. However, the residues and other structural elements that play a role in the enzyme-mediated maturation process of DENV by NS3 have yet to be definitely assigned. Identification of the binding site and the actual environment around the active site pocket are still open to questions and are of fundamental importance in understanding both the viral maturation mechanism and the significance of inhibitors nature, which may be directly responsible for future drug designing.

As a key step toward elucidating the functions of DENV NS3 and in particular for a better understanding of the active site pocket of the enzyme, the 3D structure of DEN2 NS3 serine protease to locate its key catalytic residues has been modelled. Computational comparative analyses of sequences and structures of related NS3 serine proteases have also helped to predict the electrostatic potential distribution within and around the active site pocket of NS3.

Materials and Methods

Full-length sequences of all the four serotypes of dengue virus were retrieved from NCBI genome database (*http://www.ncbi.nlm.nih.gov/*), with GenBank accession numbers DEN1 (NP_722463.1), DEN2 (NP_739587.1), DEN3 (AAM51538), DEN4 (NP_740321.1). The physiochemical properties of the NS3 sequence was carried out using Pepstats from the European Bioinformatics Institute (EBI: *http://www.ebi.ac.uk*). Protein-protein BLAST (blastp) was performed at NCBI server. Also PSI-BLAST (Protein Specific Iterated BLAST) was used to search for protein sequences that are distantly related to DEN2_NS3. This program uses an iterated search in which sequence found in one round of searching are used to build a score model for the next round of searching. Besides, pair-wise sequence alignment was performed using BLAST-2-sequences (bl2seq) at NCBI server.

Sequences were also aligned using multiple sequence alignment tool ClustalW (1.83) (*http://www.ebi.ac.uk/tools/clustalw*). This program uses a progressive method to build its alignments using the BLOSUM 62 substitution matrix for proteins. The alignments were then analysed for their pattern of conserveness. Hydrophobicity analysis was performed for DEN2 NS3 serine protease to display the distribution of polar and nonpolar residues. The plot was generated using Kyte and Doolittle scale using a program from Molecular Toolkit of Colorado State University (*http://www.cvmbs.colostate.edu/molkit/hydropathy/index.html*). Secondary structure of the protein 1BEF was predicted and analysed by using the program Hierarchical Neural Network (HNN: *http://www.expasy.org/tools/*).

Also, Ramachandran plot (*http://i.moltalk.org/*) and DeepView/Swiss-PdbViewer v3.7 (spdbv) (*http://www.expasy.org/spdbv/*) were used to model and analyse the 3D structure of 1BEF. Swiss-PdbViewer (GlaxoSmithKline R&D) is an application that allows several proteins to analyse at the same time. The proteins can be superimposed in order to deduce structural alignments and compare their active sites or any other relevant parts. Swiss-PdbViewer is tightly linked to SWISS-MODEL, an automated homology-modelling server developed by the Structural Bioinformatics Group of the Swiss Institute of Bioinformatics (SIB) at the Biozentrum in Basel.

Results and Discussion

Sequence homology analysis of NS3 serine protease from all four serotypes DENV (1-4)

Pair-wise alignment study has been carried out using the four serotypes of DENV virus. Sequence comparison showed ~78-90% identities and similarities in sequences. Comparing the 4 serotypes with the most potent DEN2 showed homology in the order of DEN2 >DEN3 >DEN1 >DEN4. In order to combat a viral infection, understanding the functional mechanism is of utmost importance. Thus to focus more on the mechanism and the structure-function relationship Multiple Sequence Alignment (MSA) study has also been carried out along with some other members of *Flavivirus* family, like West Nile virus (WNV) and hepatitis C virus (HCV) that share much homology in structure. The results showed nice gap-free conserved regions both in the amino-terminal (N-terminal) and the carboxyl-terminal (C-terminal) domains of all four serotypes. The key active site residues that are present within the N-terminal region also remain very much conserved within the four serotypes (Figure 2). Arrows indicate the position of the key active site residues that remain very much conserved within the four DEN (1-4) serotypes and in WNV-NS3. The structural comparison results also showed that in HCV-NS3 a polar uncharged Asn (N) residue is present instead of the essential negatively charged Asp (D) residue.

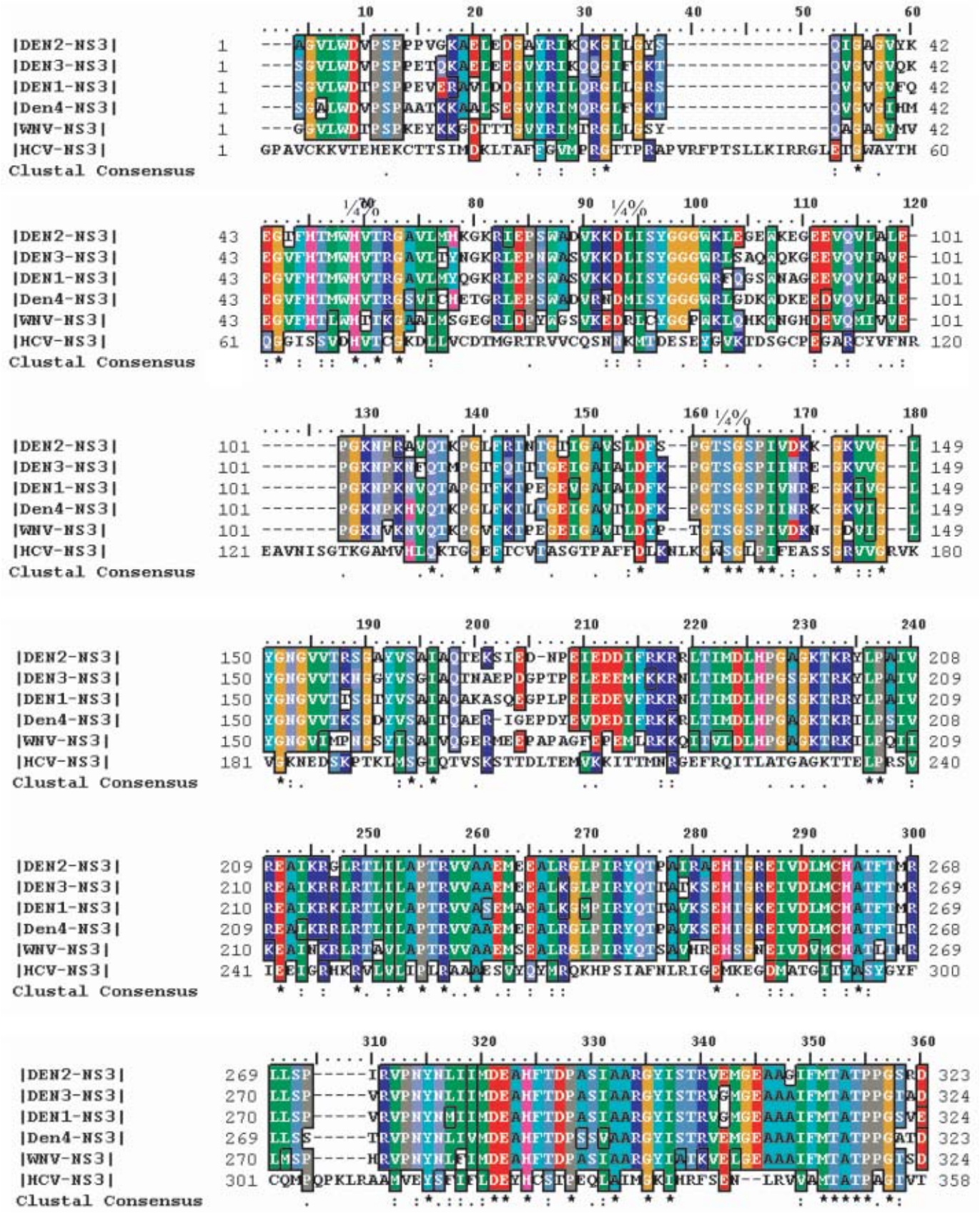


Figure 2. *Structure-based sequence alignment of the N-terminal domain of DEN2-NS3 with the NS3 proteases of the other three DENV serotypes, West Nile virus and hepatitis C virus.*

Detailed structural analysis of DEN2-NS3 serine protease

Physicochemical analysis of DEN2-NS3 serotype showed a balance in acidic and basic residues content. The presence of the number and types of amino acids like E, G, K, S, I, T, D, R are strongly responsible for destabilizing α-helical structure whereas the possibility of having a relatively extended β-sheets conformation (Table 1). The secondary structure prediction analysis further confirmed the structure. The predicted structure showed two small helices within the amino terminal region and a third helix at the carboxyl terminal end of the protein which altogether comprises ~17 % of the sequence length. More extensive β-strands structures (~28 %) and coil-coil region (~54 %) are distributed along the sequence (Figure 3). The Ramachandran plot of NS3 serine protease (Figure 4) was constructed using the crystal structure (1BEF) data further helps to predict the conformation of the backbone of the polypeptide chain more quantitatively starting from its amino acid sequences. The calculated φ(phi) and Ψ(psi) angles predicted the residues within the core region of the protein structure.

Table 1. *Pepstats analysis result showing the physiochemical properties of DEN2-NS3*

PEPSTATS of EMBOSS_001 from 1 to 181
Molecular weight = 19288.82
Residues = 181
Charge = 0.5
Average residue weight = 106.568
Isoelectric point = 6.8012
A280 molar extinction coefficient = 34850
A280 extinction coefficient 1 mg/ml = 1.81
Improbability of expression in inclusion bodies = 0.845

Residue	Number	Mole%	DayhoffStat
A = Ala	13	7.182	0.835
B = Asx	0	0.000	0.000
C = Cys	1	0.552	0.191
D = Asp	9	4.972	0.904
E = Glu	14	7.735	1.289
F = Phe	3	1.657	0.460
G = Gly	26	14.365	1.710
H = His	3	1.657	0.829
I = Ile	11	6.077	1.351
K = Lys	17	9.392	1.423
L = Leu	10	5.525	0.747
M = Met	2	1.105	0.650
N = Asn	4	2.210	0.514
P = Pro	11	6.077	1.169
Q = Gln	5	2.762	0.708
R = Arg	5	2.762	0.564
S = Ser	11	6.077	0.868
T = Thr	9	4.972	0.815
V = Val	17	9.392	1.423
W = Trp	5	2.762	2.125
X = Xaa	0	0.000	0.000
Y = Tyr	5	2.762	0.812
Z = Glx	0	0.000	0.000

Property	Residues	Number	Mole%
Tiny	(A+C+G+S+T)	60	33.149
Small	(A+B+C+D+G+N+P+S+T+V)	101	55.801
Aliphatic	(I+L+V)	38	20.994
Aromatic	(F+H+W+Y)	16	8.840
Non-polar	(A+C+F+G+I+L+M+P+V+W+Y)	104	57.459
Polar	(D+E+H+K+N+Q+R+S+T+Z)	77	42.541
Charged	(B+D+E+H+K+R+Z)	48	26.519
Basic	(H+K+R)	25	13.812
Acidic	(B+D+E+Z)	23	12.707

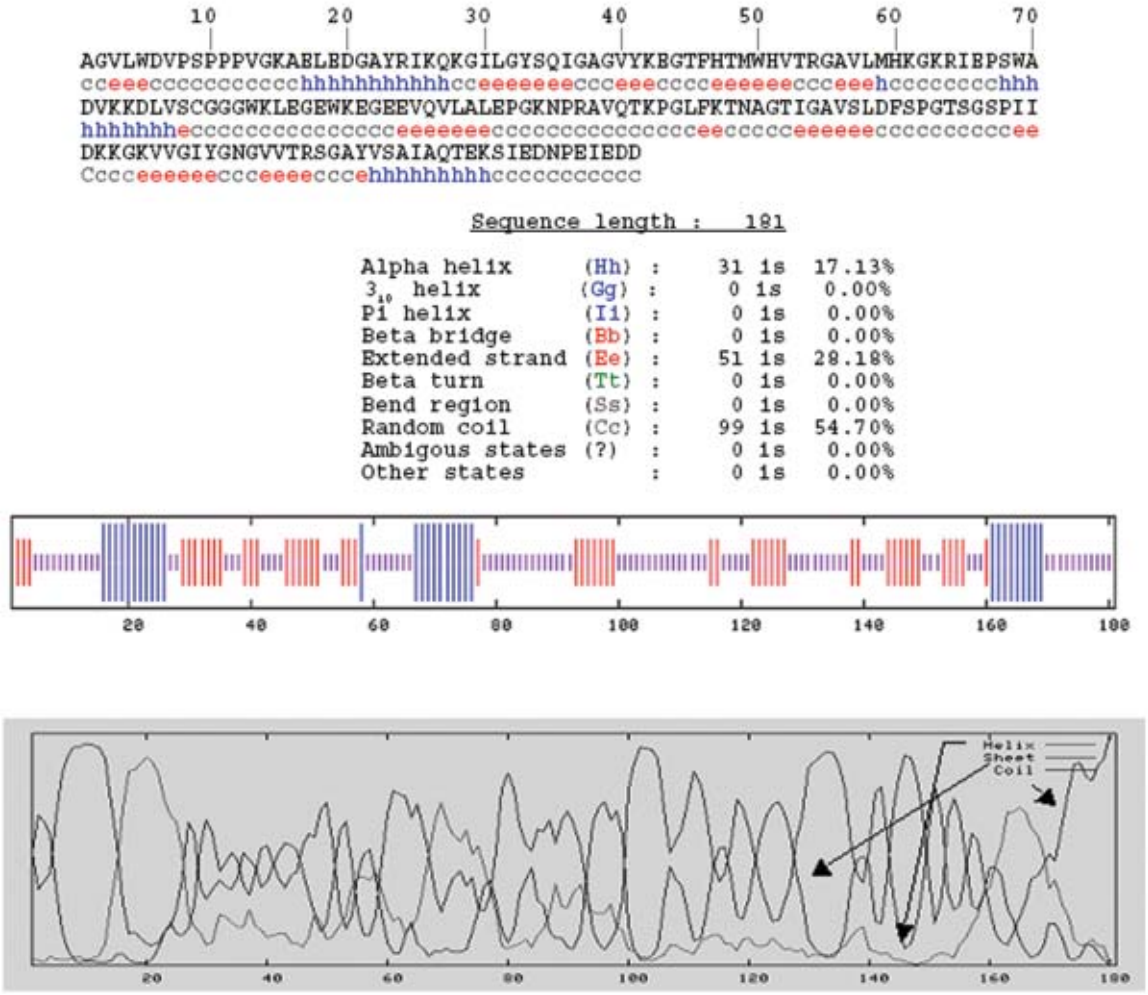
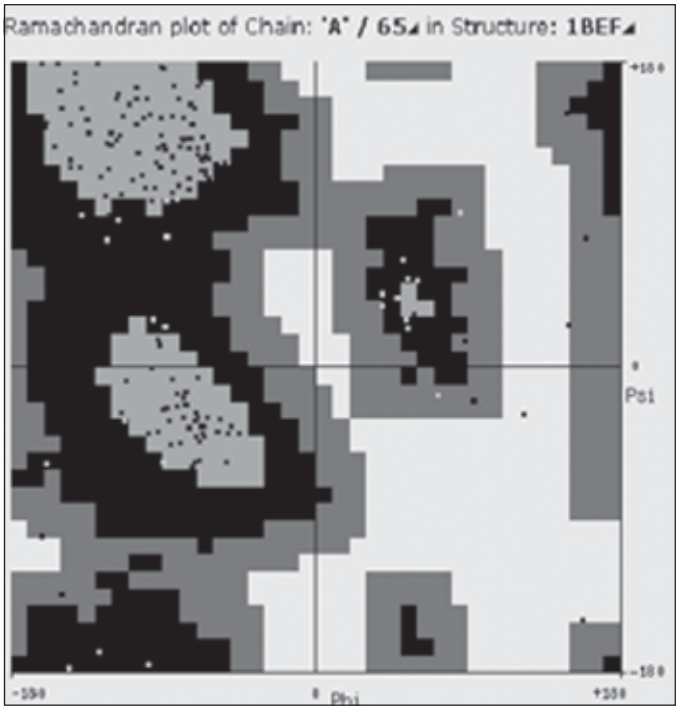


Figure 3. *Secondary structural study of DEN2-NS3 using HNN.*



The following numbers include residues with Phi/Psi angles calculated, but not GLY and PRO.	
Residues in CORE:	82.7% (115)
Residues in ALLOWED:	15.1% (21)
Residues in CORE + ALLOWED:	97.8% (136)
Residues in GENEROUS:	2.2% (3)
Residues in DISALLOWED:	0.0% (0)
Number of GLY:	25(14.3%)
Number ofPRO:	11(6.3%)

Figure 4. *Ramhachandran plot of DEN2-NS3 serine protease.*

Modelling the 3D active site pocket of DEN2-NS3 serine protease

To have a better understanding of the structure-function mechanism of NS3 serine protease of DEN2 serotype, the 3D structure of DEN2-NS3 protein was modelled and analysed using Swiss-Pdb Viewer. The crystal structure data (1BEF) along with the MSA and Ramachandran plot data were used to generate a 3D model of DEN2-NS3. Hydrogen bonds were built within the key residues of the 3D active site pocket (Figure 5A and 5B). This enables the DEN2-NS3 structure to be compared with other known serine proteases. The NS3 protease showed a chymotrypsin-like fold with β-barrels, formed by six β-strands, with the catalytic triad (His51-Asp75-Ser135) located at a cleft (Figures 5A and 5B).

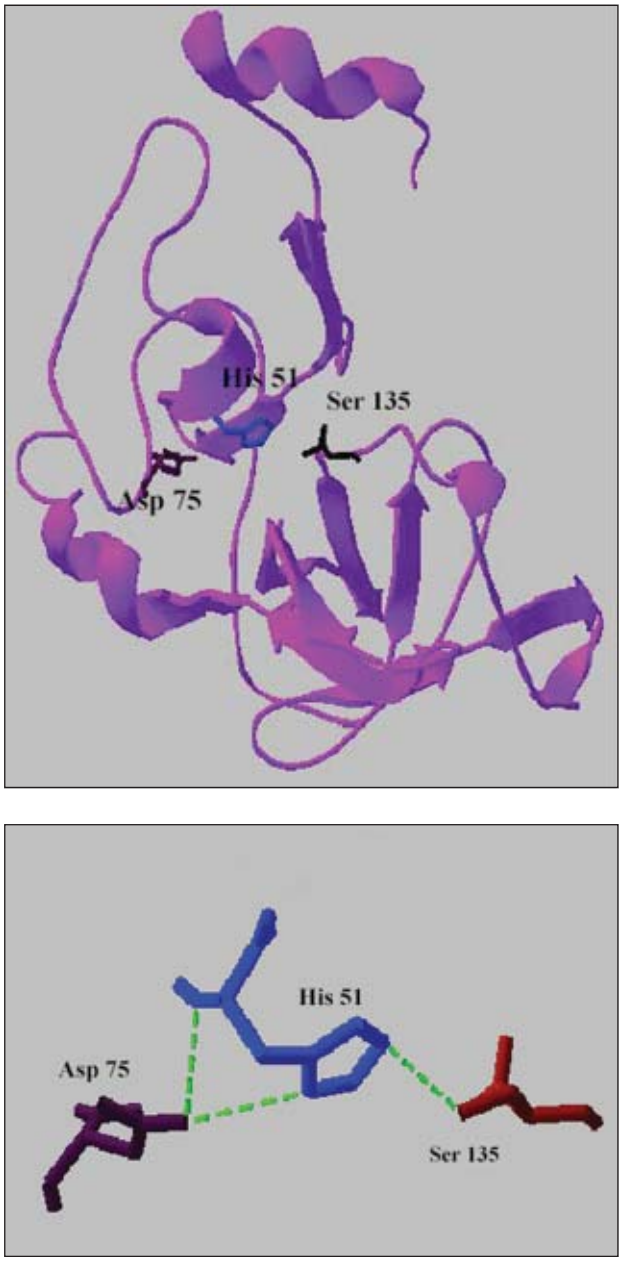


Figure 5. *Proposed model of the active site pocket of DEN2-NS3. A. Ribbon diagram of the proposed model. B. Key active site residues.*

This finding is in agreement with a recent study of NS3 protease of West Nile virus (WNV) where similar structural patterns have also been reported[14].

Analysing electrostatic potential distribution within and around the active site pocket of NS3

Hydrophobicity (or hydrophilicity) plots are designed to display the distribution of polar and nonpolar residues along a protein sequence. The goal of this analysis was to predict membrane-spanning segments (highly hydrophobic) or regions that are likely exposed on the surface of proteins (hydrophilic domains) and therefore potentially antigenic. The hydrophobic segments of DEN2-NS3 were identified using the Kyte and Doolittle scale showed quite strong hydrophobic signals within the residues (between ~75-88) on hydrophobicity scale (Figure 6). Further electrostatic potential of DEN2-NS3 serine protease was calculated using Swiss-PdbViewer. The charge distribution and electrostatic potential calculation in and around the active site pocket of NS3 serine protease gave an overall idea about the environment of the substrate-binding pocket. From this analysis, presence of two distinct regions around the active site pocket indicates a positively charged zone and another negatively charged zone (Figure 7). It is

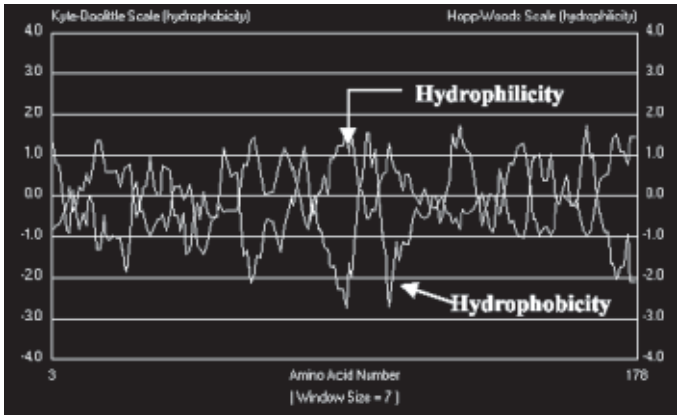


Figure 6. *A hydrophobicity analysis plot of DEN2-NS3 serine protease.*

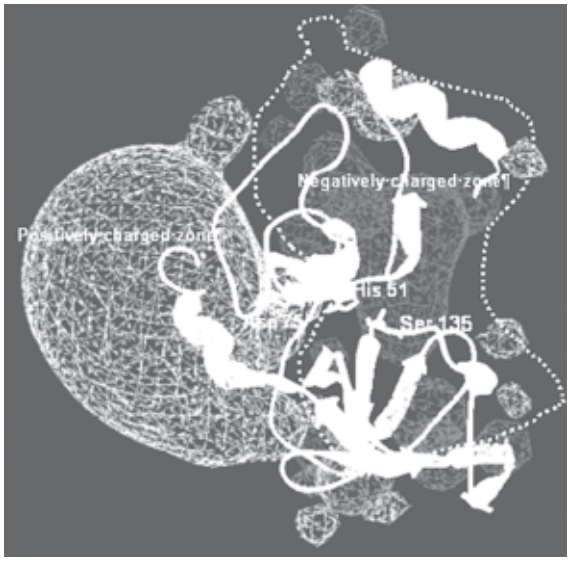


Figure 7. *Proposed electrostatic potential regions around the active site pocket of DEN2-NS3 serine protease.*

clear from the predicted potential that there is a relatively big chunk of negatively charged region than the positively charged region. These two oppositely charged chunks are located around the active site residues, facing each other, indicating that these regions might be very important for stabilizing the ES-complex.

In summary, this paper has presented a more detailed structural investigation of the NS3 serine protease of dengue virus, which reveals the common architecture shared by the four serotypes and with some other members of *Flavivirus* family, the secondary structure, the proposed model of the 3D active site pocket with key residues provides a structural framework to understand further the multifunctional properties of this enzyme. The charge distribution and electrostatic potential calculation in and around the active site pocket of DEN2-NS3 serine protease have also been predicted. This proposed model could be used for better understanding of the environment of the substrate-binding pocket. This work would, therefore, stimulate more rationally designed structure-function studies aimed at elucidating the role of this enzyme in virus maturation.

References

1. Anonymous. 2005. Division of vector-borne infectious disease. Dengue fever fact sheet. Centres for Disease Control & Prevention. Available at: *http://www.cdc.gov/ncidod/dvbid/dengueS*. Accessed September 10, 2006.
2. Rigau-Perez JG, Clark GG, Gubler DJ, Reiter P, Sanders EJ & Vorndam AV. 1998. Dengue and dengue haemorrhagic fever. *Lancet*. **352**: 971-977.
3. Chanprapaph S, Saparpakorn P, Sangma C, Niyomrattanakit P, Hannongbua S, Angsuthanasombat C & Katzenmeier G. 2005. Competitive inhibition of the dengue virus NS3 serine protease by synthetic peptides representing polyprotein cleavage sites. *Biochem Biophys Res Commun*. **330**(4): 1237-1246.
4. Murthy KHM, Clum S & Padmanabhan R. 1999. Dengue virus NS3 serine protease: Crystal structure and insights into interaction of the active site with substrates by molecular modelling and structural analysis of mutational effects. *J Biol Chem*. **274**(9): 5573-5580.
5. Chambers TJ, Weir RC, Grakoui A, McCourt DW, Bazan JF, Fletterick RJ & Rice CM. 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. *Science*. **87**(22): 8898-8902.
6. Lobigs M. 1992. Proteolytic processing of a Murray Valley encephalitis virus non-structural polyprotein segment containing the viral proteinase: Accumulation of a NS3-4A precursor which requires mature NS3 for efficient processing. *J Gen Virol*. **73**(9): 2305-2312.
7. 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**(9): 4364-4374.
8. Wengler G, Czaya G, Farber PM, Hegemann JH. 1991. *In vitro* synthesis of West Nile virus proteins indicates that the amino-terminal segment of the NS3 protein contains the active centre of the protease which cleaves the viral polyprotein after multiple basic amino acids. *J Gen Virol*. **72**(4): 851-858.
9. Bazan JF & Fletterick RJ. 1989. Detection of trypsin-like serine protease domain in flaviviruses and pestiviruses. *Virololy*. **171**(2): 637-639.
10. Gorbalenya AE, Donchenko AP, Blinov VM & Koonin EV. 1989. Cysteine protease of positive strand RNA viruse and chymotrypsin-like serine proteases: A distinct protein superfamily with a common structural fold. *FEBS Lett*. **243**(2): 103-114.
11. Chao D, King C, Wang W, Chen W, Wu H & Chang G. 2005. Strategically examining the full-genome of dengue virus type 3 in clinical isolates reveals its mutation spectra. *Virol J*. **2**: 72.
12. Lamarre D, Anderson PC, Bailey M, Beaulieu P, Bolger G, Bonneau P, Bos M, Cameron DR, Cartier M, Cordingley MG, Faucher AM, Goudreau N, Kawai SH, Kukolj G, Lagace L, LaPlante SR, Narjes H, Poupart MA, Rancourt J, Sentjens RE, St George R, Simoneau B, Steinmann G, Thibeault D, Tsantrizos YS, Weldon SM, Yong CL & Llinas-Brunet M. 2003. An NS3 Protease inhibitor with antiviral effects in humans infected with hepatitis C virus. *Nature*. **426**: 186-189.
13. Murthy HM, Judge K, DeLucas L & Padmanabhan R. 2000. Crystal structure of Dengue virus NS3 protease in complex with a Bowman-Birk inhibitor: Implications for flaviviral polyprotein processing and drug design. *J Mol Biol*. **301**: 759-767.
14. Erbel P, Schiering N, D'Arcy A, Renatus M, Kroemer M, Lim SP, Yin Z, Keller T, H, Vasudevan SG, Hommel U. 2006. Structural basis for the activation of flaviviral NS3 proteases from dengue and West Nile virus. *Nat Struct Mol Biol*. **13**(4): 372-373.