Malaysian Journal of Science **25** (1): 15 – 22 (2006)

A Revisit into the DEN2 NS2B/NS3 Virus Protease Homology Model: Structural Verification and Comparison with Crystal Structure of HCV NS3/4A and DEN2 NS3

Lee Yean Kee[1], **Rozana Othman**[1], **Habibah A Wahab**[2], **Rohana Yusof**[3] and **Noorsaadah Abd. Rahman**[1,*]

[1] Department of Chemistry, Faculty of Science, University of Malaya, 50603 Kuala Lumpur
[*] noorsaadah@um.edu.my
[2] School of Pharmaceutical Sciences, Universiti Sains Malaysia, Pulau Pinang
[3] Department of Molecular Medicine, Faculty of Medicine, University of Malaya, 50603 Kuala Lumpur.
Received 1st June 2005, accepted in revised form 27th March 2006.

ABSTRACT Although the crystal structure of DEN2 NS3 serine protease has been reported, the proteolytic mechanism of this enzyme in the presence of NS2B as co-factor which greatly enhances its activity is not well-understood. Using the homology model of DEN2 NS3 co-complexed with NS2B co-factor based on HCV NS3/4A as the proposed template, the model of the DEN2 NS2B/3 complex was reproduced and its structure evaluated through PROCHECK, VERIFY3D and ERRAT. Comparison of the homology model with the crystal structure of DEN2 NS3 revealed that this homology model served as a better choice for the NS2B/3 protein model in drug design.

ABSTRAK Walaupun struktur kristal bagi protease serina DEN2 NS3 telah dilaporkan, bagaimana mekanisme proteolitik enzim dalam kehadiran NS2B sebagai ko-faktor mempertingkatkan aktivitinya masih tidak difahami dengan mendalam. Dengan model homologi DEN2 NS3 berkompleks NS2B berasaskan templat HCV NS3/4A, model kompleks DEN2 NS2B/NS3 telah dihasilkan dan struktur dinilai menggunakan PROCHECK, VERIFY3D dan ERRAT. Perbandingan di antara model homologi dengan struktur kristal DEN2 NS3 menunjukkan model homologi adalah pilihan yang lebih tepat bagi model NS2B/3 untuk tujuan rekabentuk drug.

(Dengue, NS3, serine protease, structural verification, HCV)

INTRODUCTION

Dengue fever and dengue hemorrhagic fever are serious diseases with almost 40% of the world population currently at risk. This disease is easily spread by the *Aedes aegyptii* and *Aedes albopictus* mosquitoes which are highly populated in tropical countries. Studies revealed that about 50 million cases of dengue infection are reported every year. There is currently no available drug or vaccine to combat this disease. However, various efforts are currently being put into finding effective vaccines or inhibitors for this disease [10], [17].

The dengue fever and dengue hemorrhagic fever are caused by the dengue virus that is grouped under the flavivirus family. Currently, 4 serotypes of dengue virus have been discovered with the DEN2 serotype being the most prevalent amongst the four. The single-stranded, positive-sense RNA of dengue virus has approximately 10,723 nucleotides (for the New Guinea-C strand). The genomic RNA has a single open reading frame that encodes a polyprotein of 3,391 amino acids. These amino acids are processed into 3 structural (C, prM, and E) proteins that are further assembled into the virion and seven non-structural proteins, NS1 to NS5, and are expressed in infected cells [9]. Studies have revealed that the second largest protein encoded by the virus, NS3 to contain a serine proteinase catalytic triad within the terminal region of 180 amino acid residues which require 40 amino acid residues of NS2B for protease activity [7].

The polyprotein precursor processing occurs co-translationally as well as post-translationally and is performed by either the host signalase in association with the membranes of the endoplasmic reticulum or the viral protease. The protease cleaves the viral polyprotein at four junctions, NS2A-NS2B (Arg-Ser), NS2B-NS3 (Arg-Ala), NS3-NS4A (Lys-Ser), and NS4B-NS5 (Arg-Gly) where a pair of dibasic amino acids at the P2 and P1 positions followed by a small, non-branched amino acid at P1' was found as the consensus of substrate cleavage motif [20, and references therein].

Many approaches have been employed to understand the mechanism, structure and molecular interaction between the serine protease of NS2B/3 complex and its substrates. The minimum domain size required for protease activity of the 69-kDa NS3 protein has been mapped to 167 residues at the N terminus [16]. The virus sequence alignments analyzed revealed that structural motifs as well as the characteristic catalytic triad (His-Asp-Ser) of mammalian serine proteases are conserved in all flaviviruses [2], [8]. Sequence comparison among the serine proteases and mutational analysis verified that a catalytic triad of NS2B/3 comprised of the residues His51, Asp75, and Ser135, and that replacement of the catalytic Ser135 residue by alanine resulted in an enzymatically inactive NS3 protease [19]. The presence of a peptide co-factor is also found to be essential for optimal catalytic activity of the flaviviral proteases with natural polyprotein substrates [1], [5].

Although the dengue virus NS3 protease exhibits NS2B-independent activity with model substrates for serine proteases such as *N*-α-benzoyl-L-arginine-*p*-nitroanilide, the enzymatic cleavage of dibasic peptides is markedly enhanced in NS2B/3 complex. In addition, the presence of the NS2B co-factor has been shown to be an absolute requirement for *trans*-cleavage of a cloned polyprotein substrate [20]. A genetically engineered NS2B(H)-NS3pro protease containing a non-cleavable nonamer glycine linker between the NS2B activation sequence and the protease moiety exhibited higher specific activity with *para*-nitroanilide peptide substrates than the NS2B(H)-NS3pro molecule [15]. The NS2B-NS3pro protease incorporating a full-length NS2B cofactor sequence could catalyze the cleavage of 12-mer peptide substrates representing native polyprotein junctions [11], [12]. However, this protein appeared to be completely resistant to proteolytic self-cleavage.
A model of the NS2B/NS3 dengue virus protease was first constructed through homology modeling using the crystal structure of HCV NS3/4A complex as template by Brinkworth and co-workers, with the suggestion that the 40 amino acid residues of dengue virus NS2B co-factor could be reduced to 12 hydrophobic residues [4]. Experimental data on hepatitis C virus protease showed some structural and mechanistic explanations for the protease activation by its co-factor, where the NS4A co-factor was found to affect the folding of the NS3 protease. This resulted in conformational rearrangements of the N-terminal 28 residues of the protease and a strand displacement that lead to the formation of a well-ordered array of three β-sheets with the co-factor as an integral part of the protease fold [13], [19]. These conformational changes reorient the residues of the catalytic triad making it more favorable for proton shuttling during proteolytic process.

In the absence of the crystal structure for NS3pro complexed with its co-factor NS2B, it is rather difficult to understand the structure and conformation of the dengue virus serine protease as well as its mechanism of action. Comparative modeling could provide an alternative method to understand the structure and conformation of protease and further aid in predicting the binding interactions of the substrate with competitive substrate-based or non-substrate-based inhibitors.

Although the model of the NS2B/3 complex of the dengue virus protease has been proposed based on a homology modeling on HCV [4], the verifications of this model were not extensively carried out. In addition, there was no detailed discussion on the comparisons between the homology structure and the dengue NS3 crystal structure to enable better insights into the role of NS2B co-factor in influencing the function of NS3. Furthermore, our attempts to recrystallize the NS3 protein have been unsuccessful thus far. Hence, we aim to reproduce the dengue virus protease model through homology modeling and to subject the model for a more intensive structural verification and evaluation using server-based protein structure verification program, namely PROCHECK, VERIFY 3D and ERRAT. It is hoped that the results of these verifications will provide an insight into the problems encountered in the NS3 recrystalization.

The similarities and differences between the crystal structures and computer model of NS3 generated in this work as well as that of Brinkworth's [4] are also discussed.

MATERIALS AND METHODS

Homology model of DEN2 NS2B/3 Serine Protease

Homology model of NS2B/3 of dengue virus type 2 was built using the HCV serine protease NS3/4A (pdb ID: 1jxp) as the template. The Modeller (mod6v2) software package was used to perform model building. The sequence alignment was done based on the published results of Brinkworth *et al.* [4]. The quality of the backbone of the rough model generated from Modeller was then evaluated using PROCHECK [14], VERIFY3D [3] and ERRAT [6] on the UCLA bioinformatics server. Energy minimization (100 steps of steepest decent plus 50 steps of conjugate gradient) was performed onto the model, using Hyperchem software package (Hypercube, Inc.) to reduce the bumps and bad contacts while keeping the backbone of the protein restrained. The model evaluation was then repeated.

Comparison of the homology model with crystal structures of and DEN2 NS3 and HCV NS3/4A

The similarities and differences of the structure and conformation around the catalytic triad of the constructed homology model of DEN2 NS2B/3 serine protease were evaluated using the crystal structures of DEN2 NS3 (pdb ID: 1bef) and the HCV NS3/4A (pdb id: 1jxp).

RESULTS

Homology model building and model evaluation

The Ramachandran plot obtained from PROCHECK (Figure 1) showed an overall 100 % non-glycine residue to be in the allowed region, which implies a good protein backbone structure and folding, where the distribution of the φ/ϕ angle of the model were within the allowed region. In addition, VERIFY 3D showed 90.4% of the residues having a 3D-1D score greater than 0.2, suggesting a reasonable conformation of the residues in the model. However, region with a 3D-1D score lower than 0.2 was found in the range of Glu91-Gln110, indicating a lower confidence in its conformations and folding, implying a lower homology between DEN2 serine protease and HCV serine protease in this particular region. ERRAT was used to check the non-bonded structures of the queried structure and to compare with a database of reliable high-resolution structures. The DEN2 N2SB/3 homology model gave 77.1% overall quality factor of the sequence to be below 95% rejection limit for each chain in the input structure (see figure 1, 2 and 3 for more details) after several iterations of geometry optimizations. This indicated a better three-dimensional profile of the protein from pre-generated homology structure (data not shown). All these verification procedures performed on the NS2B/3 protease model indicated that the model has reached a satisfactory fold quality. Hence, no further loop modeling was carried out on the model.

DISCUSSION

Comparison of the homology model with crystal structures of and DEN2 NS3 and HCV NS3/4A

Overall, the homology model showed almost the same folding of protein backbone as observed in the DEN2 NS3 crystal structure where in the first domain (NH_2-terminal domain) of NS3, one alpha-helix and 6 beta sheets are found, both in the homology model as well as in the DEN2 NS3 crystal structure (Figure 4). The differences between two models, however, were observed in the second NS3 domain where more loop regions are found in the crystal structure of NS3 compared to those observed in the homology model. In addition, the crystal structure of NS3 indicated the presence of only one alpha helix and 7 beta sheets in the C terminal region, whilst one extra beta sheets in the same domain was observed in the homology model. It has been reported that in the crystal structure of NS3 with NS4A co-factor of HCV, the incorporation of NS4A co-factor into the N-terminal domain β-sheet led to a more rigid and precise framework for "prime-side" substrate binding channel residues, thus providing a better catalytic cavity that makes the NS3 enzyme active in proteolytic process [13].

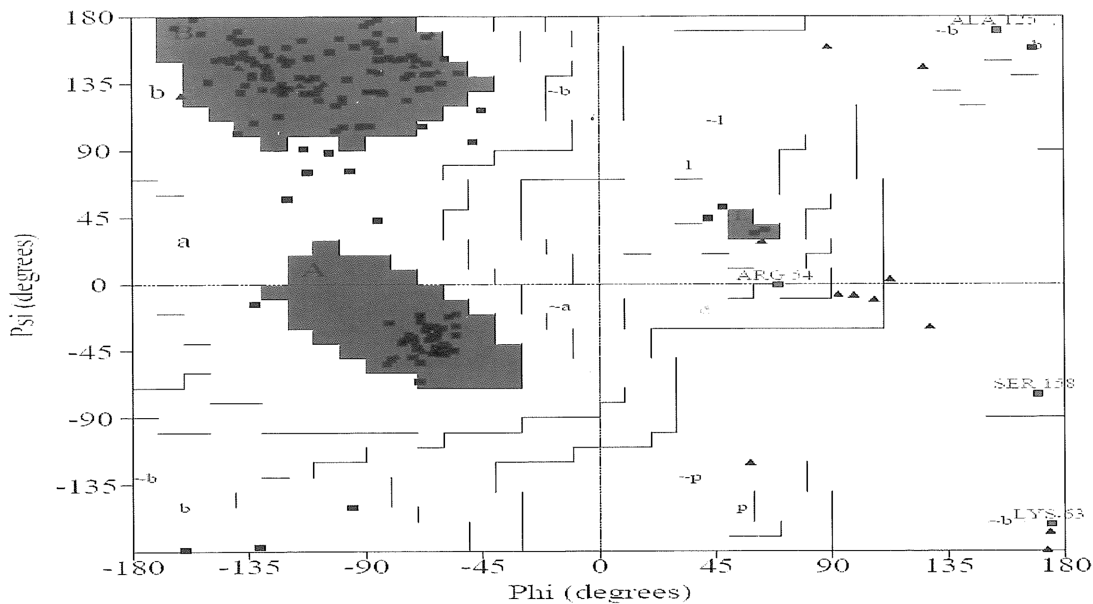

Figure 1. Ramachandran plot of built homology model of DEN2 NS2B/3 complex

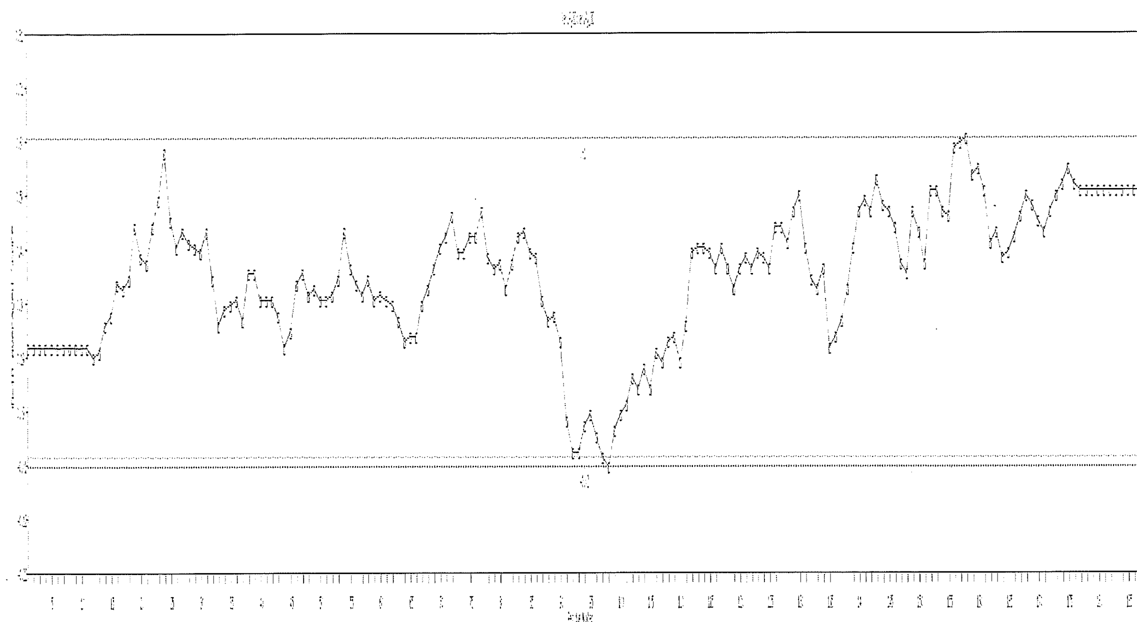

Figure 2. Verify 3D plot of NS2B/3 homology model

Superimposition (Figure 4d) of the NS3 structure from the crystallographic data with that of the homology model revealed a difference in the folding between Gly114-Val126 of crystals of NS3 compared to the homology model. These may explain the importance of the NS2B co-factor's role in re-packing the NS3 protein into a more rigid and stabilized structure, especially at the C- terminal domain, where more secondary structure was observed as compared to the NS3 crystal in the absence of the NS2B co-factor.

The catalytic triad residues for HCV NS3/4A and DEN2 NS2B/3 serine proteases were found to be in the structurally conserved region and there is no significant conformational differences observed between them. The RMSD value found between the catalytic triad residues of the HCV

NS3/4A (His57, Asp81 and Ser139) crystal with the homology model of DEN2 NS2B/3 (His51, Asp75 and Ser135) is 0.6 while that of the homology model of DEN2 NS2B/3 and the DEN2 NS3 crystal structures is 1.1. The hydrogen bonding between the hydroxyl group of Ser135 and cycloimine of His51 sidechain was observed in the catalytic triad of the reported Den2 NS3 crystal structure (Figure 5a). The side chain carboxyl oxygen of Asp75, however, is observed to be oriented away from His51 (Figure 5a). This caused it to not form a hydrogen bond between carboxyl group of Asp75 and cycloamine of His51 and hence, the proton transfer from Asp75 to Ser 135 which is required to activate the proteolytic process can not transpire.

In the homology model, the carboxyl oxygen of Asp75 and His51, as well as that of His75 and the hydroxyl of Ser135, was found to be at 1.6 Å which is within the hydrogen bonding distance (Figure 5b), suggesting a better arrangement of catalytic residues.

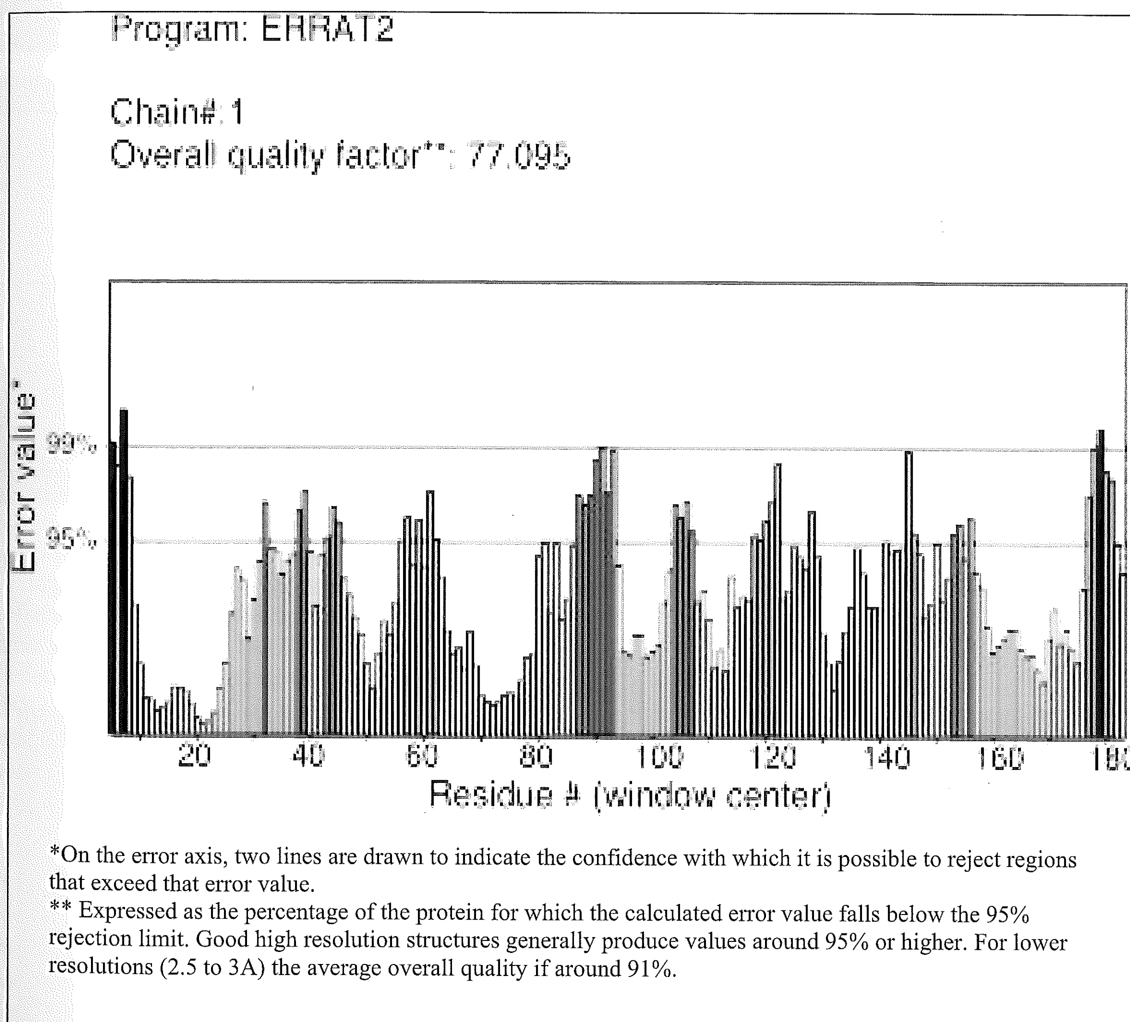

Figure 3. ERRAT analysis of NS2B/3 homology model

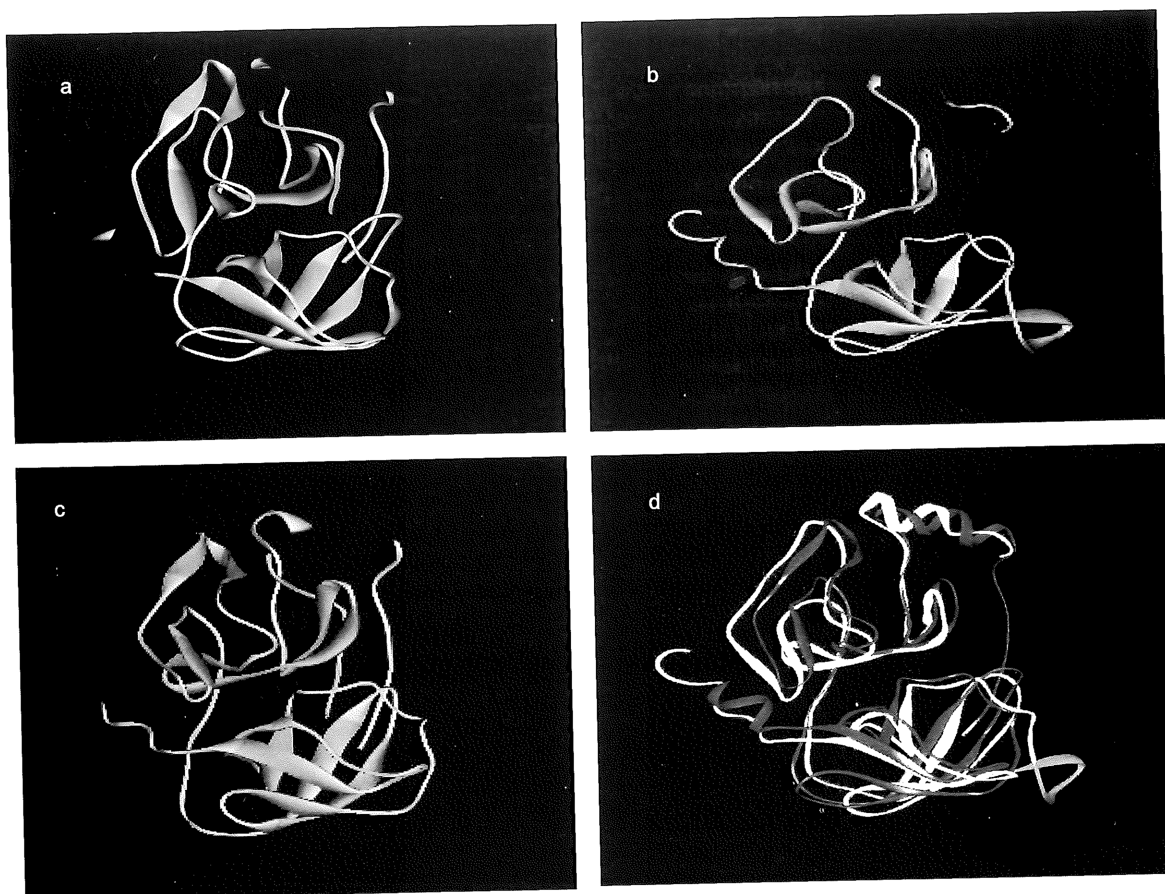

Figure 4. Structures of flavivirus serine proteases; a: DEN2 NS2B/3 complex homology model, b: DEN2 NS3 crystal structure, c: HCV NS3/4A complex crystal structure, d: superimposition of DEN2 NS3 crystal (white) and homology model (dark grey). In Figure 2d, fragment that exhibiting the difference in the protein folding is shown in lighter grey (bottom right corner of the box).

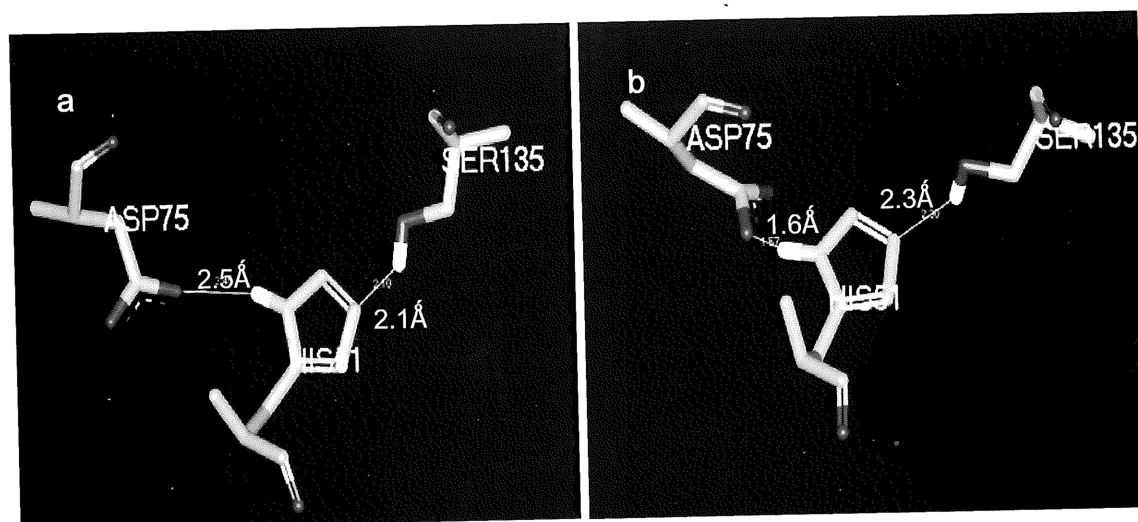

Figure 5. Spatial arrangement of catalytic triad in a: DEN2 NS3 crystal structure (pdb id: 1BEF) and b: DEN2 NS2B/3 complex homology model. Distance between the carboxyl oxygen of Asp75 and His51, as well as that of His75 and the hydroxyl of Ser135 are indicated

As demonstrated in Table 1, all the structures gave a reasonable reading of the Ramachandran plot, where more than 90% of non-glycine residues were located in the allowed region and no residues were located in disallowed region was observed in all the protein structures. This indicated that the backbone of the serine protease of HCV and DEN2 has a reasonably high degree of homology, in spite of its low sequence identity [4].

Table 1. Structural verification (PROCHECK, VERIFY3D, ERRAT) comparison between structure of HCV NS3/4A crystal, homology model of DEN2 NS2B/3 and DEN2 NS3 crystal

STRUCTURAL VERIFICATION	HCV CRYSTAL	NS2B/3 HOMOLOGY MODEL	NS3 CRYSTAL
Ramachandran Plot			
Core	80.7	86.3	82.7
Allowed	18.0	11.0	15.1
Generously allowed	1.3	2.7	2.2
Disallowed	0	0	0
VERIFY3D	97.4	90.4	52.2
ERRAT	92.0	77.1	49.4

Interestingly, the homology model for DEN2 NS2B/3 displayed a better reading in VERIFY3D and ERRAT as compared to crystal structure of NS3, suggesting a better side chain packing in the computer model. The absence of the NS2B co-factor fragment in the crystal structure of NS3 is attributed to a lower quality 3D structure of the crystals. This information has provided some insights into the role of the protease co-complexed with NS2B co-factor, which seems to re-orientate the active pocket of the DEN2 NS3, exhibiting a better side chain packing for a more efficient proteolytic cleavage [20]. However, it is still not viable to use the NS3 crystal structure as a template for generating the model structure of NS2B/NS3 dengue virus protease since the structural verification studies of various methods performed showed low confidence in the structural information. On the other hand, structural verifications performed on the crystal structure of HCV NS3/4A showed a remarkably high level of confidence. Hence, the homology model generated using HCV crystal structure as a template should provide a better and more accurate picture of the DEN2 serine protease structure.

CONCLUSION

The homology model of DEN2 NS2B/3 complex serine protease has provided information on the similarities and differences between the structures of HCV NS3/4A crystal and uncomplexed DEN2 NS3 crystal. The crystal structure of HCV NS3/4A appears to be a better choice to be used as a template in order to generate a model for DEN2 serine protease since results indicated better structure verification values for HCV NS3/4A than that of DEN2 NS3 crystal.

Acknowledgement The authors acknowledge financial support for this project provided by the Akademi Sains Malaysia under the Scientific Advancement Grant Award (SAGA-M8).

REFERENCES

1. Bartenschlager, R., V. Lohman, T. Wilkinson, and J. O. Koch. (1995). Complex formation between the NS3 serine-type proteinase of the hepatitis C virus and NS4A and its importance for polyprotein maturation. *J. Virol.* **69**: 7519–7528.
2. Bazan, J. F., and Fletterick, R. J. (1989). Detection of a trypsin-like serine protease domain in flaviviruses and pestiviruses. *Virology* **171**: 637–639
3. Bowie, J. U., Luthy, R. and Eisenberg, D. (1991). A method to identify protein sequences that fold into a known three-dimensional structure. *Science* **253** (5016): 164-170
4. Brinkworth, R. I., D. P. Fairlie, D. Leung, and P. R. Young. (1999). Homology model of the dengue 2 virus NS3 protease: Putative interactions with both substrate and NS2B cofactor. *J. Gen. Virol.* **80**: 1167–1177.

5. Chambers, T. J., Grakoui, A. and Rice, C. M. (1991). Processing of the yellow fever virus nonstructural polyprotein: A catalytically active NS3 proteinase domain and NS2B are required for cleavages at dibasic sites. *J. Virol.* **65** (11): 6042–6050.
6. Colovos, C. and Yeates, T. O. (1993). Verification of protein structures: patterns of nonbonded atomic interactions. *Protein Sci.* **2** (9): 1511-1519.
7. Falgout, B., Miller, R. H. and Lai, C.H. (1993). Deletion Analysis of Dengue Virus Type 4 Nonstructural Protein NS2B: Identification of a Domain Required for NS2B-NS3 Protease Activity. *J. Virol.* **67** (4): 2034-2042.
8. Gorbalenya, A. E., Donchenko, A. P., Koonin, E. V. and Blinov, V. M. (1989). N-terminal domains of putative helicases of flavi- and pestiviruses may be serine proteases. *Nucleic Acids Res.* **17** (10): 3889–3897.
9. Irie, K., Mohan, P. M., Sasaguri, Y., Putnak, R. and Padmanabhan, R. (1989). Sequence analysis of cloned dengue virus type 2 genome (New Guinea-C strain). *Gene* **75** (2): 197–211.
10. Kautner, I., Robinson, M. J. and Kuhnle, U. (1997). Dengue virus infection: Epidemiology, pathogenesis, clinical presentation, diagnosis, and prevention. *J. Pediatr.* **131** (4): 516–524.
11. Khumthong, R., Angsuthanasombat, C., Panyim S., and Katzenmeier, G. (2002). In vitro determination of dengue virus type 2 NS2B-NS3 protease activity with fluorescent peptide substrates. *J. Biochem. Mol. Biol.* **35** (2): 206–212.
12. Khumthong, R., P. Niyomrattanakit, S. Chanprapaph, C. Angsuthanasombat, S. Panyim and G. Katzenmeier. (2003). Steady-state cleavage kinetics for dengue virus type 2 NS2B-NS3(pro) serine protease with synthetic peptides. *Protein Peptide Lett.* **10** (1): 19–26.
13. Kim, J. L., K. A. Morgenstern, C. Lin, T. Fox, M. D. Dwyer, J. A. Landro, S. P. Chambers, Markland, W., Lepre, C. A., O'Malley, E. T., Harbeson, S. L., Rice, C. M., Murcko, M. A., Caron, P. R. and Thomson, J. A. (1996). Crystal structure of the hepatitis C virus NS3 protease domain complexed with a synthetic NS4A cofactor peptide. *Cell* **87** (2): 343–355.
14. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK: A program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* **26**: 283-291.
15. Leung, D., Schroeder, K., White, H., Fang, N. X., Stoermer, M. J., Abbenante, G., Martin, J. L., Young, P. R. and Fairlie, D. P. (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** (49): 45762–45771.
16. Li, H., Clum, S., You S., Ebner, K. E. and Padmanabhan, R. (1999). The serine protease and RNA-stimulated nucleoside triphosphatase and RNA helicase functional domains of dengue virus type 2 NS3 converge within a region of 20 amino acids. *J. Virol.* **73** (4): 3108–3116.
17. Monath, T. P. (1994). Dengue: The risk to developed and developing countries. *Proc. Natl. Acad. Sci. U. S. A.* **91** (7): 2395–2400
18. Valle, R. P. C. and Falgout, B. (1998) Mutagenesis of the NS3 protease of dengue virus type 2. *J. Virol.* **72** (1): 624–632.
19. Yan, Y., Li, Y., Munshi, S., Sardana, V., Cole, J. L., Sardana, M., Steinkuehler, C., Tomei, L., De Francesco, R., Kuo, L. C. and Chen, Z. (1998). Complex of NS3 protease and NS4A peptide of BK strain hepatitis C virus: A 2.2 Å resolution structure in a hexagonal crystal form. *Protein Sci.* **7** (4): 837–847.
20. Yusof, R., Clum, S., Wetzel, M., Murthy, H. M. and 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.