PEDS Advance Access published September 12, 2012

Protein Engineering, Design & Selection pp. 1–8, 2012
doi:10.1093/protein/gzs049

Transplantation of a hydrogen bonding network from West Nile virus protease onto Dengue-2 protease improves catalytic efficiency and sheds light on substrate specificity

Danny N.P.Doan[1†], Kun Quan Li[2†], Chandrakala Basavannacharya[1], Subhash G.Vasudevan[1,5] and M.S.Madhusudhan[2,3,4,5]

[1]Duke-NUS Graduate Medical School, Program in Emerging Infectious Diseases, 8 College Road, Singapore 169857, Singapore, [2]Bioinformatics Institute, 30 Biopolis Street, #07-01, Matrix, Singapore 138671, [3]Department of Biological Sciences, National University of Singapore, Singapore and [4]School of Biological Sciences, Nanyang Technological University, Singapore

[5]To whom correspondence should be addressed. E-mail: madhusudhan@bii.a-star.edu.sg (M.S.M); subhash.vasudevan@duke-nus.edu.sg (S.G.V)

Received February 3, 2012; revised June 21, 2012; accepted July 24, 2012

Edited by Jeffery Saven

The two-component serine protease of flaviviruses such as Dengue virus (DENV) and West Nile virus (WNV) are attractive targets for inhibitor/therapeutic design. Peptide aldehyde inhibitors that bind to the covalently tethered two-component WNV protease (WNVpro) with 50% inhibitory concentration (IC_{50}) at sub-micromolar concentrations, bind the equivalent DENV-2 protease (DEN2pro) with IC_{50} of micromolar concentrations at best. Conversely, the protease inhibitor aprotinin binds DEN2pro ~1000-fold more tightly than WNVpro. To investigate the residues that are crucial for binding specificity differences, a binding-site network of hydrogen bonds was transplanted from WNVpro onto DEN2pro. The transplantations were a combination of single, double and triple mutations involving S79D, S83N and S85Q. The mutant DENV proteases, except those involving S85Q, proved to be more efficient enzymes, as measured by their kinetic parameters. The binding affinities of the mutants to peptide inhibitors however showed only marginal improvement. Protein structure modeling suggests that the negatively charged residue cluster, Glu89–Glu92, of the NS2B cofactor may play an important role in determining substrate/inhibitor-binding specificity. These same residues may also explain why aprotinin binds more tightly to DEN2pro than WNVpro. Our results suggest that structure-based inhibitor design experiments need to explicitly consider/include this C-terminal region whose negative charge is conserved across the four DENV serotypes and also among the flavivirus family of proteases.

Keywords: Dengue virus protease/hydrogen bonding network/homology modelling/protein engineering/West Nile virus protease

Introduction

Dengue fever, caused by Dengue virus (DENV) infection, is a mosquito-borne viral disease in humans. About 2.5 billion people reside in dengue-prone areas and around 50 million are infected annually (WHO, 2009). There are no known specific anti-viral vaccines/medicines currently available against the DENV (WHO, 2009).

DENV belongs to the genus flavivirus. Other members of the genus include Murray-Valley Encephalitis virus, West Nile virus (WNV) and Yellow Fever virus. There are four serotypes of DENV (referred to as DEN1–4) and the genome encodes for three structural proteins (capsid C, membrane M and envelope E) and seven non-structural proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B and NS5). The DENV genome is translated into a poly-protein that requires post-translational processing into the individual proteins by the viral two-component NS2B-NS3pro protease and host encoded proteases (Lescar *et al.*, 2008). Despite numerous attempts to design compounds that inhibit the DENV NS2B-NS3 protease (Leung *et al.*, 2001; Chanprapaph *et al.*, 2005; Ganesh *et al.*, 2005; Yin *et al.*, 2006a,b; Ekonomiuk *et al.*, 2009; Tomlinson *et al.*, 2009), a comprehensive therapeutic agent against DENV remains elusive.

The recombinant DENV protease, $NS2B_{40}NS3_{1-185}$ (DEN2pro; and WNVpro for the WNV equivalent), consists of the NS2B hydrophilic region covalently linked by a flexible linker to the protease domain of the NS3 protein (NS3pro) (Leung *et al.*, 2001; Li *et al.*, 2005). The NS3pro is a trypsin-like serine protease with a catalytic triad consisting of His51, Asp75 and Ser135. Although the catalytic triad is wholly contained within the NS3, the functional protease requires the presence of the NS2B cofactor (Erbel *et al.*, 2006; Luo *et al.*, 2010; Phong *et al.*, 2011).

Crystal structures of the substrate-bound WNVpro reveal a closed conformation (Fig. 1a), where the NS2B wraps around NS3pro. Residues from both NS2B and NS3pro constitute the substrate-binding site (Erbel *et al.*, 2006; Aleshin *et al.*, 2007; Robin *et al.*, 2009). In contrast, crystal structures of DEN2pro are not substrate bound and show an open conformation, where the NS2B wraps the NS3 only halfway and folds back onto itself (Erbel *et al.*, 2006; Lescar *et al.*, 2008; Luo *et al.*, 2010). In the absence of the closed form of the DEN2pro structure, homology modeling can assist in functional characterization, especially around the binding sites (Wichapong *et al.*, 2010; Phong *et al.*, 2011; de la Cruz *et al.*, 2011; Knehans *et al.*, 2011). A recently reported closed form of the DEN3pro structure (Noble *et al.*, 2012) was shown to be similar to the WNVpro closed structures.

The high level of similarity (>50% overall sequence identity) notwithstanding, potential peptidic inhibitors bind more

[†]Both authors contributed equally to this work.

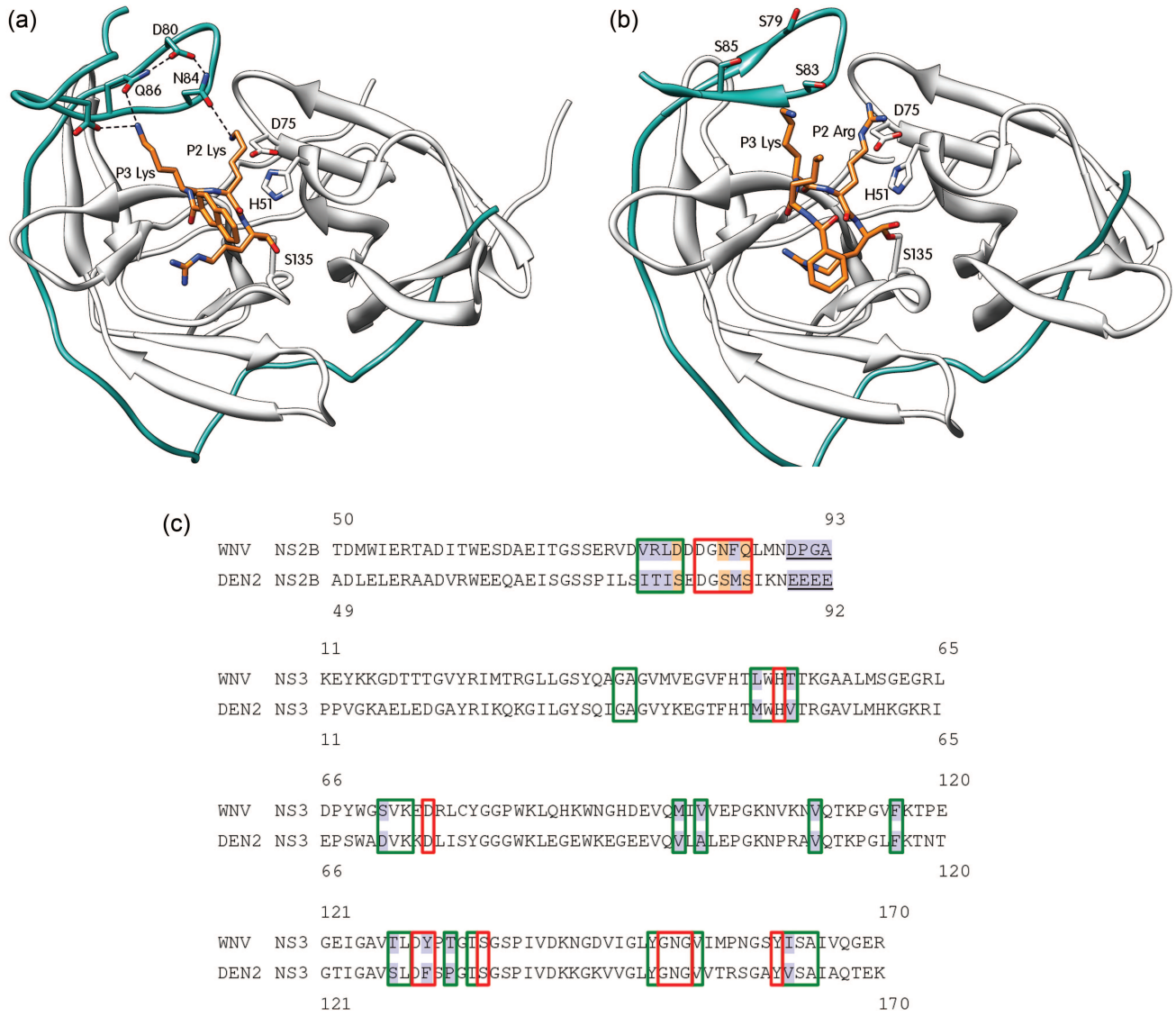

Fig. 1. Closed-form structures of WNVpro (**a**) and DEN2pro (**b**). The NS3 components are depicted as white ribbons, NS2B in blue ribbons, the peptide inhibitor, Bz-nKRR, in orange sticks. The catalytic triad, H51, D75 and S135 are also shown in stick representation. Dashed lines represent hydrogen bonds. (**c**) Sequence alignment of DEN2pro and WNVpro. The residues in Layers 1 and 2 are boxed in red and green, respectively. Non-identical residues are highlighted (blue and orange). Residue positions highlighted in orange indicate mutation sites. The 4 C-terminal residues of the NS2B construct from both proteases are underlined.

tightly to WNVpro than DEN2pro (Knox *et al.*, 2006; Yin *et al.*, 2006a,b; Stoermer *et al.*, 2008; Schüller *et al.*, 2011). This led us to explore which of the residues in the binding site of DEN2pro are significant contributors to the specificity difference. Mirroring earlier efforts to change the substrate/inhibitor binding specificity of one protease to that of another (Singh *et al.*, 2007; Shiryaev *et al.*, 2007), we endeavored to change binding site residues of DEN2pro to their WNVpro counterparts. The mutations were based on an analysis of the WNVpro crystal structures and DEN2pro homology models. We predicted that the suggested mutations would increase the affinity to the substrates/inhibitors, probably resulting in more efficient enzymatic activity, and provide insights into the roles of various binding site residues. The predictions were tested by measuring enzyme kinetics and susbtrate/inhibitor binding affinities. Homology models of DEN2pro, its mutants and WNVpro were constructed, bound to peptide inhibitors and aprotinin, to rationalize the experimental results.

The Results section first outlines the modeling efforts and the structural analysis to choose residue positions for mutagenesis studies. Experimental data from the kinetics study and the inhibition assay are then presented. With further modeling, we rationalize our predictions and its agreements and divergences from the experimental data, and the possible ramifications of our study towards designing more effective DEN2pro inhibitors.

Materials and methods

Detection of differences in binding site residues between DEN2pro and WNVpro

Stage 1: Deducing residues important for substrate/inhibitor specificity. Homology models of the closed form of DEN2pro were constructed using MODELLER 9v7 (Sali and Blundell, 1993) (Fig. 1b). The templates used were structures deposited in the protein data bank (PDB) (Berman *et al.*, 2000) of the closed form of WNVpro (PDB IDs: 2FP7

(Erbel *et al.*, 2006), 3E90 (Robin *et al.*, 2009)) and the NS3 part of the open form of DEN2pro [PDB ID: 2FOM (Erbel *et al.*, 2006)]. The three templates were first aligned structurally using the program CLICK (Nguyen *et al.*, 2011). The alignment of the sequences of the NS2B region of the target and templates was done manually, whereas the alignment of the NS3 region was retained from the structural alignment (Fig. 1c). To explore the mechanisms of ligand/inhibitor binding, the substrate-based inhibitor benzoyl-norleucine (P4)–lysine (P3)–arginine (P2)–arginine (P1)–aldehyde (Yin *et al.*, 2006b) was modeled into the DEN2pro structures. The aldehyde group was modeled as covalently bonded to the active site serine (S135).

To detect differences in the residues involved in substrate/inhibitor binding between WNVpro and DEN2pro, two layers of amino acids were considered. The first layer was constituted of residues that had at least one atom within 4 Å of any one atom of the inhibitor. The second layer was made up of residues that had at least one atomic contact within 4 Å of the first layer. Although the residues of the second layer were not directly in contact with the atoms of the inhibitor, they were deemed important in appropriately orienting the atoms of the first layer residues. Residue differences between DEN2pro (inferred from models) and WNVpro (inferred from crystal structure) in binding layers 1 and 2 are responsible for the difference in binding affinity and specificity. The site-directed mutations studied here were changes in the DEN2pro binding site to the corresponding residues from the WNVpro binding site. In all, seven mutants were created, including three single mutants (S79D, S83N and S85Q), three double mutants (S79D + S83N, S79D + S85Q, S83N + S85Q) and one triple mutant (S79D + S83N + S85Q).

Hydrogen bonds in the crystal structure and in the homology models were computed using the following criteria: 2.5 Å$<$acceptor–donor distances$<$3.5 Å and donor–acceptor–acceptor antecedent angle$>$120° (Baker and Hubbard, 1984).

Stage 2: Rationalizing experimental findings. The second set of protease–peptide inhibitor models was constructed to rationalize experimental observations and included the four glutamic acid residues in the C-terminus of the NS2B (Fig. 2). The same modeling protocol and set of templates used in stage 1 was utilized here. Models were constructed of the DEN2pro wild-type, its mutants: S79D + S83N and S85Q. Models were also constructed with the peptide inhibitor arginine (P3)–arginine (P2)–arginine (P1)–aldehyde bound to DEN2pro wild-type.

In addition to the protease–peptide inhibitor complexes, models were also constructed for DEN2pro-aprotinin and WNVpro-aprotinin complexes. Aprotinin is a 6 kDa protein that specifically inhibits DEN2pro. The WNVpro-aprotinin complex (PDB ID: 2IJO), the DEN3pro-aprotinin complex (PDB ID: 3U1J), the WNVpro structure (PDB ID: 3E90) and the DEN2pro structure (PDB ID: 2FOM) were used as templates to model the DEN2pro/WNVpro-aprotinin complexes. All aprotinin containing complexes also contained the additional four C-terminal residues of the NS2B: E89, E90, E91 and E92 for DEN2pro (Fig. 2) and D90, P91, G92 and A93 for WNVpro.

Although only one model for each system is represented in the figures (Pettersen *et al.*, 2004), all modeling exercises constructed an ensemble of 20 models using the slow refinement protocol. Inferences drawn from the models were based on averaging over the ensemble.

All homology models created for this study can be accessed at http://web.bii.a-star.edu.sg/~likq/den_hbn_2012/supplementary_materials/

Site-directed mutagenesis

Site-directed mutagenesis experiments were carried out according to the manufacturer's instruction (Quikchange Site-Directed Mutagenesis kit, Stratagene, La Jolla, CA, USA). Ten nanograms of DNA plasmid templates and 125 ng of each primer were used per reaction.

Plasmid pET15b-Den2CF40NS3pro (Li *et al.*, 2005) was used as template for the site-directed mutagenesis to generate mutation constructs S79D, S83N, S85Q and S83N + S85Q, respectively. Similarly, plasmid pET15b-Den2CF40NS3pro-S79D was used as template to generate mutation constructs S79D + S83N, S79D + S85Q and S79D + S83N + S85Q. DNA of all constructs were sequenced to verify the specific changes (1st Base, Singapore).

Protein expression and purification

Competent *Escherichia coli* BL21-CodonPlus-RIL cells (Stratagene) were transformed with pET15b-Den2CF40NS3pro expression vector and its derivatives, for subsequent protein expression and purification. The protocol for expression and purification was a slight modification of one that has been described earlier (Li *et al.*, 2005). One liter Luria-Bertani broth containing chloramphenicol (34 μg/ml) and ampicilin (100 μg/ml) was inoculated with an overnight culture and cells were grown at 37°C with shaking until OD_{600} reached ~0.6. Cell cultures were then induced with 0.4 mM isopropyl β-D-thiogalactopyranoside and allowed to grow for a further 16 h at 16°C. Cells were harvested by centrifugation at 9000 × *g* for 15 min. The cell pellet was resuspended in 20 mM sodium phosphate/500 mM NaCl/10 mM imidazole (buffer A). The cell suspension was passed through a French Press system (Thermo Electron Corporation, Milford, MA, USA) twice at 20 000 p.s.i. and the cell debris was removed by centrifugation at 30 000 × *g* for 30 min. The protein solution was filtered using a 0.22 μm filter and loaded onto a 5-ml HisTrap-FF column (GE Healthcare, USA) equilibrated with buffer A. The resin was washed with 5 column volumes of buffer A and bound proteins were eluted from the column with a linear gradient of 10 mM–500 mM imidazole in buffer A. Eluted protein fractions were analyzed by 15% sodium dodecyl sulfate-polyacrylamide gel electrophoresis and peak fractions were pooled, concentrated using Vivaspin-20 columns with 10 kDa cut-off (Sartorius Stedim, Germany) and desalted using PD-10 columns (GE Healthcare) equilibrated in 20 mM Tris-HCl pH 7/500 mM NaCl. Purified proteins were stored at −80°C in small aliquots (typically in the range of 20–60 mg/ml). The high salt concentration in the storage buffer helps stabilize the protease.

In vitro *protease assay for kinetic and inhibition study*

In vitro protease assay was performed in a final volume of 30 μl containing 50 mM Tris-HCl pH 8.5/1 mM CHAPS/20% glycerol. Briefly, each protease reaction consisted of 10 nM protein and various concentrations of the fluorogenic substrate Bz-Nle-Lys-Arg-Arg-AMC (Mimotopes, Australia). For the inhibition study, each of the three pairs of tripeptides

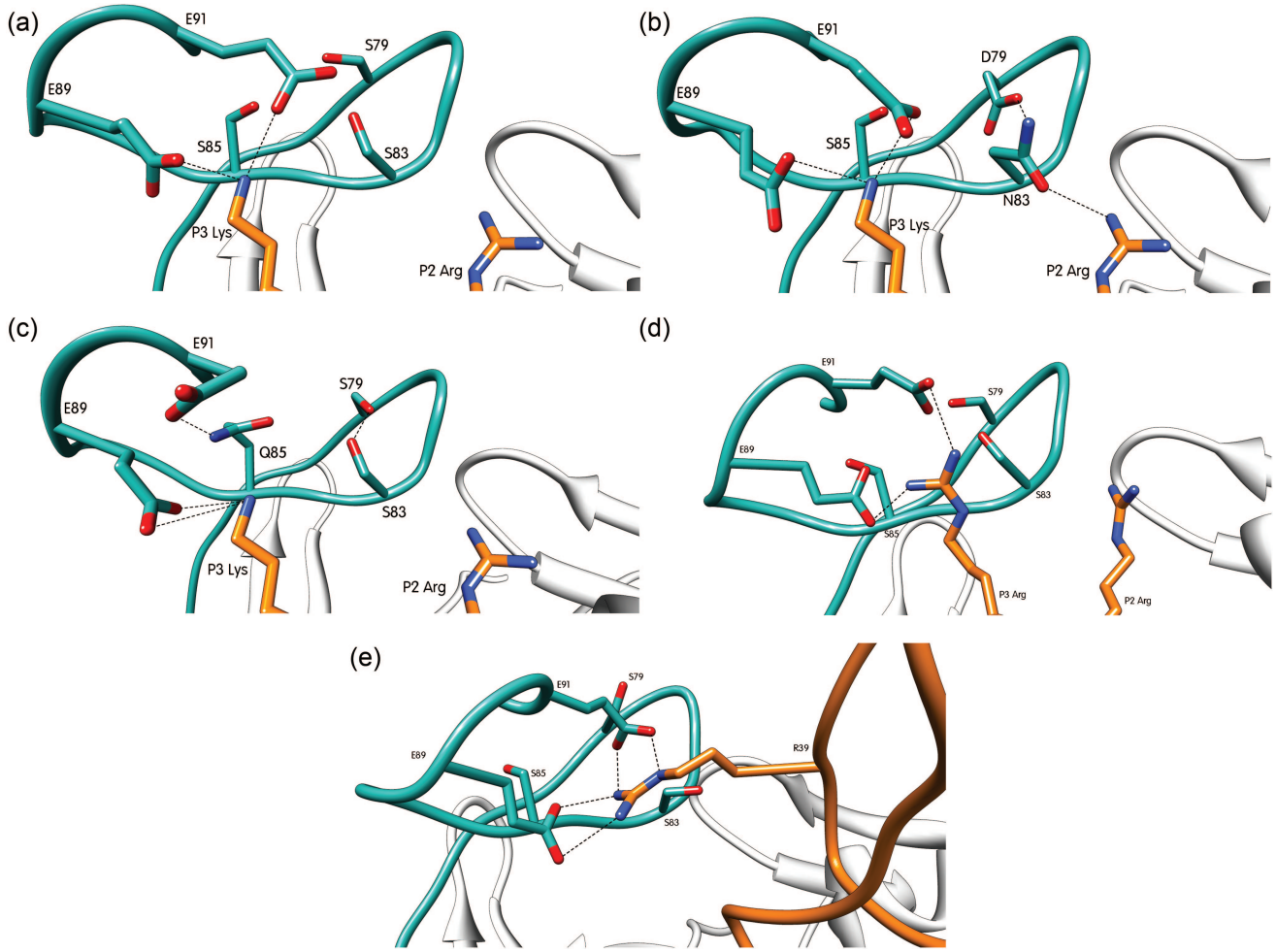

Fig. 2. Interaction of ligands (orange sticks) with the C-terminal region of the NS2B (cyan sticks) of DEN2pro (white ribbons). Hydrogen bonds (dashed lines) depicted for (**a**) DEN2pro wild-type bound to Lys–Arg–Arg, (**b**) DEN2pro mutant S79D + S83N bound to Lys–Arg–Arg, (**c**) mutant S85Q bound to Lys–Arg–Arg, (**d**) DEN2pro wild-type bound to Arg–Arg–Arg and (**e**) DEN2pro wild-type bound to aprotinin. All these models include the four glutamic acids (E89–E92) in the C-terminal region of the NS2B (cyan sticks).

with phenyl-acetyl (PheAc), bi-phenyl-acetyl (PhepheAc) or acetyl (Ac) N-terminal caps (Schüller *et al.*, 2011) were serially diluted and incubated together with 10 nM protein for 30 min at room temperature. The reaction was initiated with the addition of 20 μM of the substrate. The reaction was monitored continuously by following the increase in fluorescence at 380 nm (excitation)/450 nm (emission) using a 384-well black plate (Corning, USA) on a Tecan Infinite M200 microplate reader at 26°C. Kinetic constants were calculated using the Michaelis–Menten kinetic equation. The non-linear regression dose response curves were plotted and 50% inhibitory concentration (IC_{50}) values were determined for the inhibition study using Graphpad Prism (version 5.0, GraphPad Software, La Jolla, CA, USA). All values are reported as mean values of experiments performed in triplicate and error bars indicate the standard deviation.

Results

Homology modeling reveals binding site differences between DEN2pro and WNVpro

The homology models of DEN2pro were compared with the crystal structure of WNVpro to determine the difference in the substrate/inhibitor binding site. Thirty-nine residues were designated as belonging to the binding site. These residues were categorized into two classes: layers 1 and 2. The 14 residues that constituted layer 1 were deduced from the model to be in direct contact with the inhibitor. Layer 2 consisted of 25 residues that were in contact with the layer 1 residues but not directly in contact with the atoms of the inhibitor (see materials and methods for details). In all, 10 of the 14 residues in layer 1 were identical (Fig. 1c).

In order to transplant the binding specificity of WNVpro onto DEN2pro, the layer 1 residues should be almost identical. The only different layer 1 residue from the NS3 is Y130 in WNVpro that is F130 in DEN2pro, which we considered as a conservative substitution. The other three different layer 1 residues occur in the NS2B region. The equivalent of the sequence stretch NFQ (residues 84–86) in WNVpro is the sequence stretch SMS (residues 83–85) in DEN2pro. The residues F85 and M84 from WNVpro and DEN2pro, respectively, were deemed not to be important in specificity, as their side chains are oriented away from the substrate/inhibitor. The other two residues are involved in a hydrogen bonding network in WNVpro.

The hydrogen bonding network involving N84 and Q86 of WNVpro also included D80 from layer 2. D80 coordinated the positioning of side chains of N84 and Q86. The Oδ1 and

Oδ2 atoms of D80 are acceptors to the Nδ2 and Nε2 atoms of N84 and Q86, respectively. As a consequence, the Oδ1 and Oε2 atoms of N84 and Q86 are oriented towards the P3 and P2 residues of the inhibitor with which they can form hydrogen bonds. The crystal structures (PDB ID: 2FP7 and 3E90) suggest that these interactions are crucial for the interactions between the enzyme and the inhibitor (Fig. 1a). The residues occupying equivalent positions in the DEN2 NS2B are S79, S83 and S85. No equivalent hydrogen bond network is formed in DEN2 NS2B (Fig. 1b). These three residues from DEN2 NS2B were chosen for further mutational studies in an effort to transplant this hydrogen bond network in DEN2pro.

Binding kinetics of DEN2pro binding site mutants

The catalytic rate (k_{cat} values) of the DEN2pro, its mutants and the WNVpro were compared with one another (Table I). Of the three single mutants, only S83N exhibited similar catalytic activity as the wild-type protein (k_{cat} $\sim$0.6 s^{-1}), while S79D and S85Q mutants showed decreased ($\sim$50%) catalytic activity (k_{cat} $\sim$0.3 s^{-1}) (Table I). The double mutants S79D + S83N showed slightly increased catalytic activity (k_{cat} $\sim$0.7 s^{-1}), while the other two double mutants that involved S85Q showed decreased catalytic activity (k_{cat} $\sim$0.2 s^{-1}). The catalytic activity of the triple mutant (k_{cat} of $\sim$0.9 s^{-1}) is almost 1.5 times that of wild-type DEN2pro.

The K_m values of S79D and S85Q, the single mutants with decreased catalytic activity, showed markedly different behavior. While S79D bound the substrate more tightly, substrate affinity as implied by the K_m value was much lower for the S85Q mutant. As a consequence, the specificity constant (k_{cat}/K_m), the steady-state kinetic parameter that reflects enzyme efficiency, of the S85Q mutant was almost 3–4 times lower compared with DEN2pro and WNVpro.

Interestingly, the double mutant S79D + S83N and the triple mutant are enzymatically more efficient than both DEN2pro and WNVpro. Both mutant proteins have relatively high k_{cat} values and low K_m values (Table I). While the K_m value of the triple mutant is close to that of the WNVpro K_m value (133.9 μM and 148.3 μM respectively), it has a larger k_{cat} value ($\sim$0.9 s^{-1} compared with $\sim$0.5 s^{-1}). The triple mutant is 2 times and 3 times more efficient than WNVpro and DEN2pro, respectively. Of all the wild-type and mutants enzymes, the triple mutant has the greatest catalytic rate with a k_{cat} value of $\sim$0.9 s^{-1}. The S79D + S83N double mutant has the smallest K_m value of 65 μM, making its affinity 3 times and 4 times stronger than WNVpro and DEN2pro, respectively. In contrast, the S85Q mutation resulted in the most inefficient enzyme with a specificity constant of $\sim$800 M^{-1} s^{-1}, which is only a third of the specificity constant of wild-type DEN2pro.

Table I. Kinetic constants of DEN2pro wild-type, DEN2pro mutants and WNVpro wild-type

Protease	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m (M^{-1} s^{-1})
DEN2pro	0.64 ± 0.04	270 ± 30	2400 ± 100
S79D	0.31 ± 0.02	100 ± 20	3000 ± 300
S83N	0.67 ± 0.05	240 ± 30	2800 ± 200
S85Q	0.36 ± 0.04	460 ± 70	780 ± 40
S79DS83N	0.74 ± 0.03	70 ± 10	11300 ± 1100
S79DS85Q	0.21 ± 0.01	210 ± 20	1000 ± 50
S83NS85Q	0.22 ± 0.02	260 ± 50	800 ± 90
S79DS83NS85Q	0.92 ± 0.06	130 ± 20	6900 ± 600
WNVpro	0.50 ± 0.06	150 ± 40	3400 ± 400

Inhibiting enzymatic activity with tripeptides

The inhibition of the wild-type and mutant proteins was tested using six different peptide-aldehydes. The tripeptides Lys–Arg–Arg or Lys–Lys–Arg were each capped at their N-termini with one of three head groups, phenyl-acetyl (PheAc), bi-phenyl-acetyl (PhePheAc) or acetyl (Ac) (Schüller *et al.*, 2011). Each inhibitor terminates with an aldehyde warhead. These peptide aldehydes were previously tested against both DEN2pro and WNVpro (Schüller *et al.*, 2011). The IC_{50} values of these tripeptides with WNVpro were in the range of 50 nM to $\sim$1 mM. In contrast, the IC_{50} values for DEN2pro wild-type were in the high micro-molar range with a very clear preference for the Lys–Arg–Arg peptide aldehyde over Lys–Lys–Arg. For all the six tri-peptide inhibitors used, a general trend emerged in the ranking of IC_{50} values (Table II): S85Q > S83N + S85Q > S79D + S85Q > wild-type DEN2pro > S83N > S79D > S79D + S83N + S85Q > S79D + S83N > wild-type WNVpro. The IC_{50} values of DEN2pro mutants were higher by an order of magnitude in comparison to the wild-type WNVpro IC_{50} values. The three residues mutated in this study did not significantly improve the binding affinity of the inhibitors to DEN2pro.

Modeling the C-terminal region of the NS2B

To explain anomalous kinetics of the double mutant S79D + S83N and the single mutant S85Q, models of the DEN2pro, wild-type and mutants in complex with the inhibitor were created to include the four glutamates at the C-terminus of $NS2B_{CF40}$ (Fig. 2a–c). In the single mutant, Q85 forms a hydrogen bond with E91 (Fig. 2c) and prevents it from directly interacting with the P3 Lys of the inhibitor. This accounts for the higher K_m value in comparison to the wild-type DEN2pro (Fig. 2a and Table I).

In the S79D + S83N double mutant, the hydrogen bond network is only partially transplanted. The models show that N83 forms a hydrogen bond with P2 Arg of the inhibitor. The P3 Lys of the inhibitor, as in the wild-type DEN2pro complex, can hydrogen bond with the E89 and E91 of the NS2B C-terminal region. Together, the transplanted hydrogen bonds reduce the K_m value for the mutant (Table I).

Although some of the mutants show increased enzymatic efficiency (Table I), this does not result in significantly tighter inhibitor binding (Table II). To understand this, models of DEN2pro-aprotinin (Fig. 2e) and WNVpro-aprotinin complexes (results not shown; models deposited in supplementary information) were built. From these models, it is apparent that for optimal interaction between potential inhibitors and DEN2pro, a network of hydrogen bonds similar to that between R39 of aprotinin and E89 and E91 of the NS2B needs to be formed. Replacing the P3 Lys with an Arg does not reproduce the aprotinin complex like hydrogen bond network (Fig. 2d). Hence the binding of DEN2pro to the inhibitor is bereft of a hydrogen bonding network similar to that found in the DEN2pro-aprotinin complex or even the WNVpro-inhibitor complex. It is interesting to note that the tight hydrogen bonding network found in the DEN2pro-aprotinin complex is not reproduced in the WNVpro-aprotinin

Table II. IC50 values for different tri-peptide inhibitors tested against DEN2pro wild-type, DEN2pro mutants and WNVpro wild-type

Protease	PheAc-KRR-H (μM)	PheAc-KKR-H (μM)	PhePheAc-KRR-H (μM)	PhePheAc-KKR-H (μM)	Ac-KRR-H (μM)	Ac-KKR-H (μM)
DEN2pro	7.86 ± 0.47	122.60 ± 12.87	13.08 ± 0.68	7.91 ± 0.37	55.86 ± 0.81	220.35 ± 36.27
S79D	3.40 ± 0.23	88.96 ± 4.27	5.80 ± 0.01	4.20 ± 0.28	22.96 ± 0.81	53.17 ± 2.00
S83N	5.12 ± 0.10	63.76 ± 1.09	10.92 ± 0.43	3.52 ± 0.22	38.17 ± 0.57	42.23 ± 1.38
S85Q	18.34 ± 1.21	337.40 ± 34.51	33.96 ± 0.12	22.74 ± 0.3	102.45 ± 1.63	338.30 ± 77.64
S79DS83N	2.15 ± 0.08	61.43 ± 5.62	3.52 ± 0.36	3.36 ± 0.05	10.77 ± 1.39	44.97 ± 1.31
S79DS85Q	9.66 ± 0.33	265.50 ± 6.65	18.06 ± 0.64	11.49 ± 0.58	50.50 ± 1.44	100.50 ± 2.96
S83NS85Q	16.17 ± 0.98	122.95 ± 2.47	22.33 ± 0.39	7.09 ± 0.52	94.81 ± 2.63	83.82 ± 4.92
S79DS83NS85Q	2.53 ± 0.14	52.20 ± 0.31	5.75 ± 0.71	3.35 ± 0.39	18.81 ± 0.89	41.87 ± 3.46
WNVpro	0.06 ± 0.01	0.68 ± 0.01	0.98 ± 0.10	0.05 ± 0.00	3.05 ± 0.36	0.65 ± 0.04

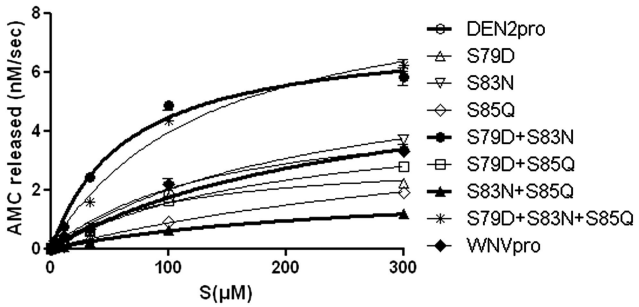

Fig. 3. Michaelis–Menten curves for WNVpro, DEN2pro and its mutants. The kinetic constants are calculated using the equation $v = V\max [S]/(K_m + [S])$.

complex. This explains why aprotinin binds more tightly to DEN2pro (~30 nM) than it does to WNVpro (low μM) (Aleshin *et al*., 2007; Mueller *et al*., 2007).

Discussions

In an effort to glean the contribution of different binding site residues towards substrate/inhibitor specificity, we transplanted a hydrogen bond network found only in WNVpro onto DEN2pro (Fig. 2a). In the WNV NS2B, this hydrogen bond network involves coordination of the side chains of critical substrate binding residues, N84 and Q86, by D80. Alanine mutations of D80, N84 and Q86 in WNV NS2B decrease protease activity to 1, 12 and 46% of the WNVpro, respectively (Chappell *et al*., 2008). We believe that this is due to the disruption of the hydrogen bond network formed between these residues. Also, the k_{cat}/K_m values of N84A and Q86A WNVpro mutants were 1 and 5% of the wild-type protease, respectively (Chappell *et al*., 2006). These data suggest that N84 and Q86 are involved in binding the substrate P2 and P3 residues in the WNVpro (Chappell *et al*., 2008; Knehans *et al*., 2011). The equivalent residues in DEN2pro, S83 and S85, were hence deduced to be important in substrate binding (Knehans *et al*., 2011). Thus, in theory, S79D, S83N and S85Q mutations should transplant the WNVpro H-bond network onto the DEN2pro, increasing its binding affinity towards substrates and peptide inhibitors.

From the homology models of the DEN2pro mutants, it was predicted that the order of binding affinity to the inhibitors would follow the trend (from high to low affinity)—WNVpro > triple mutant > double mutants > single mutants > DEN2pro. This trend was generally observed in the binding affinity study albeit with only a nominal increase in binding affinities (Table II).

The general trend notwithstanding, some of the experimental observations were surprises. The k_{cat}/K_m value of the S79D + S83N mutant was found to be 4 times larger as compared with the DEN2pro and 3 times larger as compared with the WNVpro (Table I and Fig. 3). The triple mutant also shows a large increase in k_{cat}/K_m value compared with both DEN2pro (~3 times) and WNVpro (~2 times). These increases in k_{cat}/K_m values indicate that the transplantation of the WNVpro residues into DEN2pro is beneficial to substrate binding (Fig. 2b). It is interesting to note that the addition of a third mutation, S85Q, to the S79D + S83N double mutant actually decreases the enzymatic activity.

Homology models with an extended C-terminal NS2B help in explaining these surprises. In our initial homology models (Knehans *et al*., 2011), the four C-terminal glutamates were omitted. Most, if not all, previous structural studies also do not consider this sequence stretch. The only available template to model this region was one of the two molecules in the asymmetric unit of one WNVpro crystal structure (PDB: 3E90 chains A and B). The C-terminal end of NS2B now contains four glutamate residues (E89–E92). Models of DEN2pro and its mutants show that the inhibitor interacts with the negatively charged C-terminal region (Fig. 2a–c) via hydrogen bonds and salt bridges. These interactions appear to be the strongest in the case of the double mutant S79D + S83N. The hydrogen bond network involving D79 and N83 holds the P2 residue of the inhibitor in place, while the P3 residues is hydrogen bonded to the side chains of E89 and E91 (Fig. 2b). The presence of a substantial negative charge in this C-terminal region enhances the stability of the P3 interaction. We conjecture that this is the reason why the double mutant, S79D + S83N, shows the best catalytic efficiency.

The S85Q mutation, whether alone or as a part of a double mutant, was always associated with decreased relative binding affinity in the DEN2pro (Tables I and II). Residues in the negatively charged C-terminal interact with the mutated (S85Q) residue (Fig. 2c). A consequence of this interaction is that the hydrogen bond between Q85 and the P3 Lys of the inhibitor is disrupted. This is also probably why the transplantation of the entire hydrogen bond network from WNVpro onto DEN2pro is worse (in terms of catalytic efficiency) than the double mutant, S79D + S83N, which does not contain the mutation S85Q.

These terminal Glu residues all lie within the first layer of the binding site (Fig. 1c). Further, the charge in this region

Flavivirus Proteases	NS2B C Terminal Region
Dengue Virus Type 1	K^{87}DEERDDT^{94}
Dengue Virus Type 2	K^{87}NEEEEQT^{94}
Dengue Virus Type 3	K^{87}DDETENI^{94}
Dengue Virus Type 4	R^{87}DVEETNM^{94}
Murray Valley Encephalitis Virus	L^{88}NDPGVPW^{95}
Japanese Encephalitis Virus	I^{88}DDPGVPW^{95}
West Nile Virus	M^{88}NDPGAPW^{95}
Yellow Fever Virus	L^{87}SEEKVPW^{94}

Fig. 4. Sequence alignment of the NS2B C-terminal region in flavivirus proteases. The negatively charged residues are highlighted. A colour version of this figure is available as supplementary data at *PEDS* online.

of NS2B is conserved among flaviviruses (Fig. 4). As seen in the WNVpro crystal structure (PDB: 3E90), D90, a negatively charged residue, in this region stabilizes the binding of the inhibitor via a salt bridge. Our models of DEN2pro also suggest a similar interaction between the E89 and the P3 Lys of the inhibitor. The presence of three other glutamates in the immediate neighborhood of E89 (especially E91) leads us to infer that these negatively charged residues play a direct and important role in binding to the positively charged P3 residue via electrostatic interactions. This is further borne out in our models of the DEN2pro-aprotinin complex. The aprotinin, via the charged residues K15 and R39, forms electrostatic interactions with the DEN2pro (Fig. 2e). The R39 in the aprotinin is located such that all three of its guanidinium nitrogen atoms can hydrogen bond with the negatively charged C-terminal region of the NS2B. In contrast, the interaction of the arginine at the P3 position of the peptide inhibitor with this C-terminal region is weaker and does not involve the Nε atom in any hydrogen bonding (Fig. 2d). The C-terminus of the NS2B of WNVpro is different from that of DEN2pro in that it does not have this succession of four glutamates. While this explains why the aprotinin binds more tightly to DEN2pro, it clearly shows the importance of the C-terminal region in inhibitor binding and possibly determining binding specificity.

Crystal structures of the closed form of DEN3pro were solved recently (Noble *et al.*, 2012; PDB ID: 3U1I). A comparison of the crystal structure of our models of DEN2pro (65% sequence identity to DEN3pro) showed remarkable similarity (C^{α} root mean square deviation (RMSD) of 0.79 Å over 194 residues). These crystal structures do not contain the negatively charged C-terminal region of NS2B. As there is no observed charge–charge interaction between the NS2B and the P3 Lys side chain, the authors concluded that the charged P3 residue could be replaced with an uncharged residue. This is contrary to other DEN2pro studies that observed a preference for the positively charged Lys at P3 position of the ligand (Li *et al.*, 2005; Stoermer *et al.*, 2008). Our models strongly suggest that the four residues in the C-terminal region, particularly E89 and E91, determine substrate/inhibitor-binding specificity. Moreover, these residues would be part of layer 1 of the inhibitor interacting residues. The inclusion of these residues in structural studies would improve efforts in designing potential inhibitors, as they are likely to be the most important determinants in substrate/inhibitor specificity.

Supplementary data

Supplementary data are available at *PEDS* online.

Acknowledgements

We thank Prof Yin Zheng, Nankai University and Drs Brian Chia and Jeffrey Hill of Experimental Therapeutics Centre, Singapore for tripeptide inhibitors used in this study.

Funding

This work is in part supported by DUKE-NUS Signature Research Program (funded by the Agency for Science, Technology and Research, Singapore and the Ministry of Health, Singapore), and the National Medical Research Council, Singapore (www.nmrc.gov.sg) under grant NMRC/GMS/051/2011 to S.G.V.

References

Aleshin,A.E., Shiryaev,S.A., Strongin,A.Y. and Liddington,R.C. (2007) *Protein Sci.*, **16**, 795–806.
Baker,E.N. and Hubbard,R.E. (1984) *Prog. Biophys. Mol. Biol.*, **44**, 97–179.
Berman,H.M., Westbrook,J., Feng,Z., Gilliland,G., Bhat,T.N., Weissig,H., Shindyalov,I.N. and Bourne,P.E. (2000) *Nucleic Acids Res.*, **28**, 235–242.
Chanprapaph,S., Saparpakorn,P., Sangma,C., Niyomrattanakit,P., Hannongbua,S., Angsuthanasombat,C. and Katzenmeier,G. (2005) *Biochem. BioPhys. Res. Commun.*, **330**, 1237–1246.
Chappell,K.J., Stoermer,M.J., Fairlie,D.P. and Young,P.R. (2006) *J. Biol. Chem.*, **281**, 38448–38458.
Chappell,K.J., Stoermer,M.J., Fairlie,D.P. and Young,P.R. (2008) *J. Gen. Virol.*, **89**, 1010–1014.
de la Cruz,L., Nguyen,T.H., Ozawa,K., Shin,J., Graham,B., Huber,T. and Otting,G. (2011) *J Am. Chem. Soc.*, **133**, 19205–19215.
Ekonomiuk,D., Su,X., Ozawa,K., Bodenreider,C., Lim,S.P., Otting,G., Huang,D. and Caflisch,A. (2009) *J. Med. Chem.*, **52**, 4860–4868.
Erbel,P., Schiering,N., D'rcy,A., *et al.* (2006) *Nat. Struct. Mol. Biol.*, **13**, 372–373.
Ganesh,V.K., Muller,N., Judge,K., Luan,C., Padmanabhan,R. and Murthy,K.H.M. (2005) *Bioorg. Med. Chem.*, **13**, 257–264.
Knehans,T., Schüller,A., Doan,D.N., Nacro,K., Hill,J., Güntert,P., Madhusudhan,M.S., Weil,T. and Vasudevan,S.G. (2011) *J. Comput. Aided Mol. Des.*, **25**, 263–274.
Knox,J.E., Ma,N.L., Yin,Z., *et al.* (2006) *J. Med. Chem.*, **49**, 6585–6590.
Lescar,J., Luo,D., Xu,T., Sampath,A., Lim,S.P., Canard,B. and Vasudevan,S.G. (2008) *Antiviral Res.*, **80**, 94–101.
Leung,D., Schroder,K., White,H., Fang,N., Stoermer,M.J., Abbenante,G., Martin,J.L., Young,P.R. and Fairlie,D.P. (2001) *J. Biol. Chem.*, **276**, 45762–45771.
Li,J., Lim,S.P., Beer,D., *et al.* (2005) *J. Biol. Chem.*, **280**, 28766–28774.
Luo,D., Wei,N., Doan,D.N., Paradkar,P.N., Chong,Y., Davidson,A.D., Kotaka,M., Lescar,J. and Vasudevan,S.G. (2010) *J. Biol. Chem.*, **285**, 18817–18827.
Mueller,N.H., Yon,C., Ganesh,V.K. and Padmanabhan,R. (2007) *Int. J. Biochem. Cell Biol.*, **39**, 606–614.
Nguyen,M.N., Tan,K.P. and Madhusudhan,M.S. (2011) *Nucleic Acids Res.*, **39**, 24–28.
Noble,C.G., She,C.C., Chao,A.T. and Shi,P.Y. (2012) *J. Virol.*, **86**, 438–446.
Pettersen,E.F., Goddard,T.D., Huang,C.C., Couch,G.S., Greenblatt,D.M., Meng,E.C. and Ferrin,T.E. (2004) *J. Comput. Chem.*, **25**, 1605–1612.
Phong,W.Y., Moreland,N.J., Lim,S.P., Wen,D., Paradkar,P.N. and Vasudevan,S.G. (2011) *Biosci. Rep.*, **31**, 399–409.
Robin,G., Chappell,K., Stoermer,M.J., Hu,S., Young,P.R., Fairlie,D.P. and Martin,J.L. (2009) *J. Mol. Biol.*, **385**, 1568–1577.
Sali,A. and Blundell,T.L. (1993) *J. Mol. Biol.*, **234**, 779–815.
Schüller,A., Yin,Z., Brian Chia,C.S., Doan,D.N., Kim,H.K., Shang,L., Loh,T.P., Hill,J. and Vasudevan,S.G. (2011) *Antiviral Res.*, **92**, 96–101.

Shiryaev,S.A., Ratnikov,B.I., Aleshin,A.E., Kozlov,I.A., Nelson,N.A., Lebl,M., Smith,J.W., Liddington,R.C. and Strongin,A.Y. (2007) *J. Virol.*, **81**, 4501–4509.
Singh,A., Walker,K.J., Sijwali,P.S., Lau,A.L. and Rosenthal,P.J. (2007) *Protein Eng. Des. Sel.*, **20**, 171–177.
Stoermer,M.J., Chappell,K.J., Liebscher,S., Jensen,C.M., Gan,C.H., Gupta,P.K., Xu,W., Young,P.R. and Fairlie,D.P. (2008) *J. Med. Chem.*, **51**, 5714–5721.
Tomlinson,S.M., Malmstrom,R.D., Russo,A., Mueller,N., Pang,Y. and Watowich,S.J. (2009) *Antiviral Res.*, **82**, 110–114.
WHO. (2009) Dengue and dengue haemorrhagic fever. Factsheet No 117.
Wichapong,K., Pianwanit,S., Sippl,W. and Kokpol,S. (2010) *J. Mol. Recognit.*, **23**, 283–300.
Yin,Z., Patel,S.J., Wang,W., *et al*. (2006a) *Bioorg. Med. Chem. Lett.*, **16**,*nbsp;369.
Yin,Z., Patel,S.J., Wang,W., *et al*. (2006b) *Bioorg. Med. Chem. Lett.*, **16**,*nbsp;403.