Journal of General Virology (1999), **80**, 1167–1177. Printed in Great Britain

Homology model of the dengue 2 virus NS3 protease: putative interactions with both substrate and NS2B cofactor

Ross I. Brinkworth,[1] David P. Fairlie,[1] Donmienne Leung[1] and Paul R. Young[2,3]

[1,2] Centre for Drug Design and Development[1] and Department of Microbiology[2], University of Queensland, St Lucia, Brisbane 4072, Australia

[3] Sir Albert Sakzewski Virus Research Centre, The Royal Children's Hospital, Herston, Brisbane 4029, Australia

The crystal structure coordinates of the hepatitis C virus NS3 protease (HCVpro) were used to develop an homology model of the dengue 2 virus NS3 protease (DEN2pro). The amino acid sequence of DEN2pro accommodates the same α-helices, β-sheets and protein-binding domains as its HCVpro counterpart, but the model predicts a number of significant differences for DEN2pro and its interactions with substrates and cofactor. Whereas HCVpro contains a Zn^{2+}-binding site, there is no equivalent metal-binding motif in DEN2pro. It is possible that the structural role played by the zinc ion may be provided by a salt bridge between Glu^{93} and Lys^{145}. The two-component viral protease comprises NS3 and a virus-encoded cofactor, NS4A for HCV and NS2B for DEN2. Previous studies have identified a central 40 amino acid cofactor domain of the dengue virus NS2B that is required for protease activity. Modelling of the putative interactions between DEN2pro and its cofactor suggests that a 12 amino acid hydrophobic region within this sequence (^{70}GSSPILSITISE^{81}) may associate directly with NS3. Modelling also suggests that the substrate binds in an extended conformation to the solvent-exposed surface of the protease, with a P1-binding site that is significantly different from its HCV counterpart. The model described in this study not only reveals unique features of the flavivirus protease but also provides a structural basis for both cofactor and substrate binding that should prove useful in the early design and development of inhibitors.

Introduction

Dengue viruses infect up to 20 million people annually and are a significant cause of morbidity and mortality in tropical and sub-tropical regions throughout the world (World Health Organization, 1996). Like other flaviviruses, the dengue viruses are enveloped and contain a single, positive-sense RNA genome of about 11 kb that encodes a large polyprotein precursor. Co- and post-translational proteolytic processing gives rise to three structural and at least seven non-structural proteins, in the gene order NH_2-C-prM-E-NS1-NS2A-NS2B-NS3-NS4A-NS4B-NS5-COOH (Fig. 1; Chambers *et al.*, 1990*a*). Flavivirus replication is dependent on correct cleavage of this polypeptide and requires both host cell proteases and a virus-encoded, two-component protease, NS2B–NS3 (Falgout *et al.*, 1991). NS2B–NS3 has been shown to cleave in *cis* at the NS2A/NS2B and NS2B/NS3 junctions and in *trans* at the NS3/NS4A and NS4B/NS5 sites (Chambers *et al.*, 1990*b*; Preugschat *et al.*, 1990, 1991). In addition, NS2B–NS3 also mediates cleavage within the viral proteins C, NS3 and NS4A (Lobigs, 1993; Arias *et al.*, 1993; Teo & Wright, 1997; Lin *et al.*, 1993). Proteolytic processing of the remaining sites is carried out by a number of host proteases (Fig. 1). Signal peptidase cleaves at the C/prM, prM/E, E/NS1 and NS4A/NS4B junctions (Speight *et al.*, 1988; Nowak *et al.*, 1989), while an unknown host protease, shown to reside in the endoplasmic reticulum, is responsible for cleavage of the NS1/NS2A site (Falgout & Markoff, 1995). The cleavage of prM to generate the M protein is mediated late in virion maturation by furin in a post-Golgi acidic compartment (Stadler *et al.*, 1997).

The presence of a trypsin-like serine protease within the N-terminal one-third of the flavivirus NS3 protein was first proposed by Bazan & Fletterick (1989, 1990) and Gorbalenya *et al.* (1989). Their analysis of virus sequence alignments revealed that structural motifs as well as the characteristic

Author for correspondence: Paul Young.
Fax +61 7 32531401. e-mail p.young@mailbox.uq.edu.au

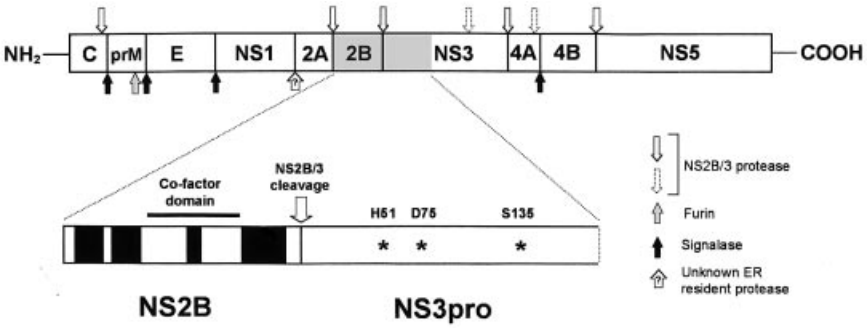


Fig. 1. Flavivirus polyprotein processing. The upper schematic shows cleavage sites on the polyprotein cleaved by host-encoded proteases (arrows below) and the virus-encoded protease complex NS2B–NS3 (arrows above). Secondary cleavage at sites within individual proteins are shown as dotted arrows. The NS2B cofactor and the proteolytic domain of NS3 (NS3pro) are shaded, with the lower schematic highlighting key features. These include the catalytic triad of NS3 (asterisks), conserved hydrophobic domains of NS2B (solid boxes) and a 40 amino acid hydrophilic region (black bar) shown to be the minimal requirement for functional NS2B cofactor involvement in the catalytic activity of NS3.

catalytic triad (His–Asp–Ser) of mammalian serine proteases were conserved in all flaviviruses. Subsequent biochemical studies confirmed the location of the protease activity within the N-terminal 180 residues of NS3 (Chambers *et al.*, 1990*b*; Preugschat *et al.*, 1990) and site-directed mutagenesis experiments verified the critical requirement for the predicted catalytic triad residues (Chambers *et al.*, 1990*b*; Wengler *et al.*, 1991). Co-expression studies demonstrated that the proteolytic activity of NS3 appeared to be dependent on the presence of NS2B (Falgout *et al.*, 1991), with the essential domain responsible for this cofactor activity being localized to a central, 40 amino acid hydrophilic sequence (Falgout *et al.*, 1993). The consensus substrate-cleavage motif for this two-component protease consists of a pair of dibasic amino acids at the P2 and P1 positions followed by a small, non-branched amino acid at P1′ (Chambers *et al.*, 1990*a*).

Crystal structures of serine proteases encoded by members of both the *Togaviridae* (Choi *et al.*, 1991, 1997) and the *Flaviviridae* (Love *et al.*, 1996; Kim *et al.*, 1996; Yan *et al.*, 1998) have been reported. However, only the crystal structure of the NS3 protease of hepatitis C virus (HCV) (Kim *et al.*, 1996; Yan *et al.*, 1998) incorporates a virus-encoded cofactor, as found in the flaviviruses. The basic features of that protease are two six-stranded β-barrel domains, characteristic of the chymotrypsin-like fold, that are separated by a linker region. In addition, there is an intimately bound structural peptide encoded by NS4A that functions as the cofactor-equivalent of NS2B in the flaviviruses. The potential of using these structural data to extrapolate to proteases encoded by other members of the *Flaviviridae* was recently addressed (Ryan *et al.*, 1998). Our goal in this study was to develop a structural model of the dengue 2 virus NS3 protease (DEN2pro) based on both sequence and structural homology with the HCV NS3 protease (HCVpro) crystal structure, in order to gain early insight into the structure and function of the protease and its likely binding to substrates and cofactors and thus facilitate substrate-based inhibitor design. The derived model has allowed us to identify: (i) a 12-residue domain of NS2B that probably constitutes the main cofactor interaction with NS3; (ii) a possible salt bridge in DEN2pro that substitutes for the structure-stabilizing zinc-binding motif; and (iii) putative interactions between DEN2pro and its substrates.

Methods

The sequence of DEN2pro (strain NGC) used in these studies was derived from clones generated and sequenced in this laboratory. Other sequences used in the study were: HCV strain BK (Yan *et al.*, 1998), dengue virus [type 1, Fu *et al.* (1992); type 3, Osatomi & Sumiyoshi (1990); type 4, Mackow *et al.* (1987)], West Nile virus (WNV) (Castle *et al.*, 1986), Japanese encephalitis virus (JEV) (Sumiyoshi *et al.*, 1987) and yellow fever virus (YFV) (Rice *et al.*, 1985).

The model of DEN2pro described in this study uses the crystal structure of HCVpro (Yan *et al.*, 1998) as a template. The atomic coordinates were obtained from the pdb file 1jxp, representing the crystal structure of the complex between HCVpro and a synthetic oligopeptide corresponding to residues Gly^{21}–Arg^{34} of the NS4A cofactor.

The alignment of sequences was optimized manually to produce a maximum number of matches of core aliphatic and aromatic residues (Trp, Phe, Tyr) in the regions of putative secondary structure, along with conserved glycine and proline residues and the regions surrounding the catalytic histidine, aspartate and serine residues (Fig. 2). Where gaps were present in the sequences, these were restricted to the loop regions. Gaps were kept to a minimum consistent with the other criteria.

The Homology module of the Insight II software package (Molecular Simulations) was used for the homology modelling based on the optimized alignment. The aligned sequences were segregated into structurally conserved regions (SCRs) and loops. SCRs are defined as regions of secondary structure, namely α-helices and β-sheets. The Homology program copies the coordinates of the backbone atoms and β-carbons from the SCRs of the template protein to the protein being modelled and new side-chains are added. Where the loop regions in the two proteins have the same number of residues, the coordinates of the modelled protein are produced in the same way. In five loop locations, C0/D0, B1/C1, A2/B2, B2/C2 and C2/helix α_b, the two proteases differed in the number of residues, and therefore the coordinates of

Table 1. Homology modelling: structural loops inserted in selected regions of the HCVpro crystal structure used for modelling of DEN2pro

DEN2pro loop region	(Tail) Loop* (Tail)	Pdb file*	Deviation†/ RMSD‡
α_0/D0	(^{14}GKAEL) EDGA (YRIKQ^{27})		
	(^{124}LQTAV) TYQP (VSVAL^{137})	2act	1·38/1·38
B1/C1	(^{40}VYKE) GTFH (TMWHV^{52})		
	(^{150}VLKI) GEHT (PSALA^{162})	1ald	1·00/1·51
A2/B2	(^{98}LALEP) GKNP (RAVQT^{111})		
	(^{473}PEIYV) PEQD (TFYKQ^{486})	1gof	0·30/0·64
B2/C2	(^{109}VQTKP) GLFKTN (AGTI^{123})		
	(^{88}GAITH) TGASGN (NFVG^{102})	1fus	0·45/0·91
C2/α_b	(^{120}AGTI) GAVS (LDFS^{131})		
	(^{343}FHVH) QFGD (NTQG^{354})	1cob	0·85/1·67

* Obtained from the Brookhaven database.
† Deviation provides a measure of the similarity in the distances between the open ends of the loops.
‡ RMSD refers to the match in conformation between the tails of the selected loop and those of the model.

DEN2pro at these locations could not be copied directly from the HCVpro structure. The B1/C1 loop is particularly short. In these cases, the Homology function Loop Search was used to locate suitable loop structures from the Brookhaven database that could be used to splice into the model. The criteria for loop selection were the correct number of residues in the loop, the distance between the ends, and the secondary structure adjacent to both ends. In each case, ten loop structures were presented and the final choice was made on the best fit, as assessed predominantly by the deviation and root-mean-squared deviation (RMSD) values. Deviation refers to how closely the loop matches the distance between the open ends. RMSD provides a measure of how well the conformation of the 'tails' (the five residues either side of the loop) of the selected loop match the 'tails' on the model being constructed. In both cases, the smaller the number the better, and it should be $< 2{\cdot}0$. The loop splice function is used to repair the omega angles and distances of where the loops are joined, to 180 ° and 1·34 Å, respectively.

The loop selections at the five locations are shown in Table 1. Following the initial determination of the coordinates of DEN2pro plus NS2B cofactor, the model was energy-minimized, employing the Discover module, using steepest descents and conjugate gradients. These minimizations produce a model with correct bond lengths and bond angles and where individual atoms are not too close together. During the initial minimizations, the backbone atoms of the SCRs of DEN2pro were kept fixed. Loop atoms were not fixed, and the final conformations of the loops are usually altered significantly from their original structures as found by Loop Search or homology modelling.

The putative cofactor domain of NS2B was modelled on the structure of the HCV NS4A cofactor as a template. Only 12 residues of the cofactor are resolved in the HCV NS3–NS4A complex crystal structure and the putative 12 residue equivalent of NS2B was determined on the basis of sequence alignment and conserved hydrophobic character, as discussed below.

To assess the validity of the model, we performed a Profiles 3D analysis (Lüthy *et al.*, 1992) using Insight II software, which provides an estimate of the number of misfolded residues. For the enzyme–cofactor complex only nine misfolded residues were identified. The analysis yielded a value of 62·5, well within the acceptable range of 37·5–83·4 for a protein of equivalent size, indicating the reliability of the model.

Results and Discussion

Structural model of DEN2pro

The HCV protease crystallized by Yan *et al.* (1998) constitutes the N-terminal 179 residues of NS3. The corresponding sequence of DEN2pro modelled in this paper consists of the N-terminal 175 residues (Fig. 2). Secondary structure elements (α-helices and β-strands) are represented schematically and labelled according to chymotrypsin notation. Identical residues are shown in bold. Overall identity between the two sequences is only 14·8%, but this figure is misleadingly low. The regions surrounding the putative catalytic residues or 'homology boxes' (boxes 1–3 in Fig. 2), as defined by Bazan & Fletterick (1989), along with some of the β-strands, particularly βE2 (box 4), which is involved in substrate interactions, show a high level of identity. In other regions, notably the B1 and B2 strands, the inter-domain linker and the C-terminal helix, sequence identity is particularly low. The derived homology model may be less accurate in these locations. Fig. 2 also shows the NS3pro sequences of a selection of other flaviviruses, with a level of residue conservation, particularly in regions of secondary structure, which suggests that the model may be broadly applicable to flavivirus serine proteases.

The amino acid sequence of DEN2pro can accommodate the same α-helices and β-sheets as HCVpro (Fig. 2), with a three-dimensional structure comprising two protein domains that adopt a chymotrypsin-like fold (Fig. 3). The N-terminal domain 1 (residues 31–79) is a β-barrel structure consisting of six β-strands, βA1–βF1, and one α-helix, α_a, which lies between βC1 and βD1. Domain 2 (residues 97–161) is also a β-barrel consisting of six β-stands, βA2–βF2 plus a helix, α_b, located between βC2 and βD2. There is also an N-terminal extension to

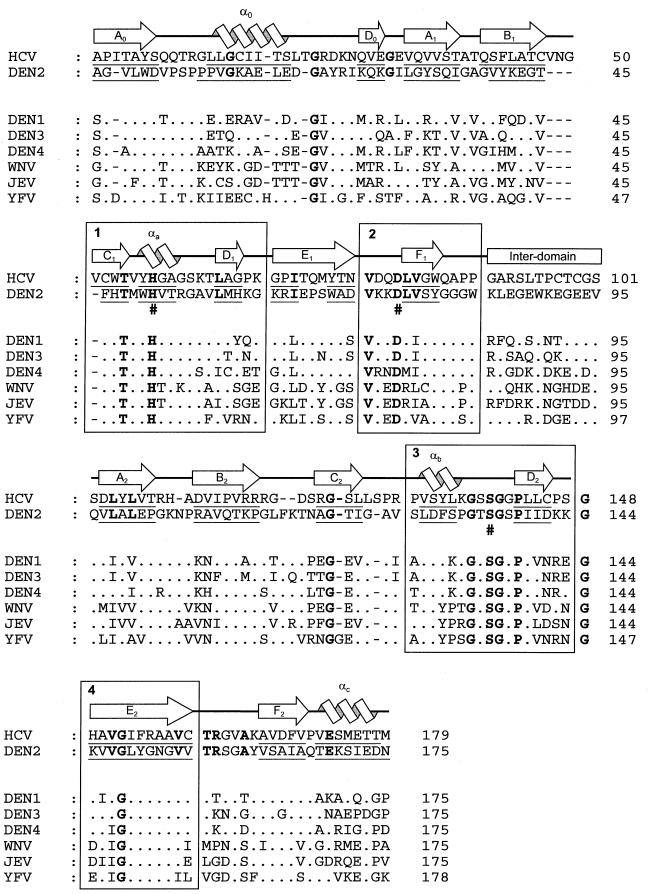


Fig. 2. Sequence alignment of NS3pro encoded by HCV, DEN2 and other flaviviruses. The catalytic triad residues are denoted by hashes (#). Secondary structure elements, labelled in accordance with chymotrypsin notation, are underlined and represented schematically above the alignment. Residues conserved between the HCV and DEN2 sequences are shown in bold, as are those residues completely conserved amongst the other flaviviruses. Residues in common with the DEN2 sequence are shown as dots, with gaps introduced into the sequence to optimize alignment as dashes. The boxes labelled 1–4 identify regions of significant similarity surrounding the catalytic triad residues (boxes 1–3) and residues likely to be involved in substrate binding (boxes 3 and 4), as originally proposed by Bazan & Fletterick (1989).

domain 1 consisting of two β-strands, βA0 and βD0, plus a helix, α_0 (residues 1–30), which extends the domain 1 β-barrel, an inter-domain linker (residues 80–96) and a C-terminal helix, α_c (residues 162–175). The residues that constitute the catalytic triad of HCVpro (His^{57}, Asp^{81} and Ser^{139}) have equivalents in DEN2pro in His^{51}, Asp^{75} and Ser^{135}. Although these residues are located in different protein domains (His^{51} and Asp^{75} in domain 1 and Ser^{135} in domain 2), the triad comes together along a cleft formed at the interface between the two domains (Fig. 3).

One striking difference between HCVpro and the flavivirus NS3 protease is the structural association of a zinc ion in the former. A metal-binding motif in HCVpro was first proposed by De Francesco *et al.* (1996) and subsequently confirmed in the crystal structures (Love *et al.*, 1996; Kim *et al.*, 1996; Yan *et al.*, 1998). Zinc binds to HCVpro via tetrahedral coordination to four ligands, Cys^{97}, Cys^{99}, Cys^{145} and a water molecule within hydrogen-bonding distance of the His^{149} side-chain (Fig. 4 *b*; Love *et al.*, 1996). These residues are located in the loop region separating the two domains (Cys^{97} and Cys^{99}) and in the βD2–βE2 turn (Cys^{145} and His^{149}). The zinc ion is too far away from the active site of HCVpro to do anything other than stabilize the structure by holding the two protease domains together. This structural feature is similar to the role played by disulphide bridges in many other serine proteinases (e.g. Cys^{136}–Cys^{201} in elastase), which also anchor the βD2–βE2 turn to the inter-barrel loop (Love *et al.*, 1996). For DEN2pro it would appear that, instead of a zinc-binding site, there may be a salt bridge (Fig. 4 *a*) consisting of Glu^{93} and Lys^{145} (instead of Cys^{99} and His^{149} in HCVpro). Such a salt bridge may provide structural stability and also explain the metal-independence of enzyme activity for DEN2pro. Significantly, all dengue virus protease sequences contain these same two charged residues. These residues are not completely conserved in all flaviviruses however, and for these there may be alternative structure-stabilizing interactions within this region.

Interactions between DEN2pro and NS2B cofactor

A cofactor is required for enzyme activity of HCVpro and this is supplied by downstream residues in NS4A (Bartenschlager *et al.*, 1994). Biochemical, mutagenesis and peptide-based studies have revealed that a relatively short, central hydrophobic region of NS4A (Gly^{21}–Lys^{34}; see upper panel in Fig. 5 *a*) is all that is required for NS3pro binding and protease activation (Lin *et al.*, 1995; Butkiewicz *et al.*, 1996). The ability of a short synthetic peptide to provide the complete cofactor requirement in *trans* (Lin *et al.*, 1995) prompted its inclusion in the crystal structure determination of the catalytically active HCV protease complex, NS3pro–NS4 [Gly^{21}–Pro^{39} in Kim *et al.* (1996); Gly^{21}–Arg^{34} in Yan *et al.* (1998)]. These studies showed that the NS4A peptide forms an extended β-strand that is sandwiched between βA0–α_0 and βA1 of NS3pro (see Fig. 3). Indeed, the N-terminal domain of NS3pro acts as a sort of flap to bury the cofactor sequences within a hydrophobic core. As a consequence, the cofactor is probably an integral structural component of the complex and almost certainly plays a significant role in the overall stabilization of NS3pro (see below). All of the contacts with NS3pro span the region encompassing residues Gly^{21}–Ser^{32} of NS4A (Kim *et al.*, 1996), with all but two of the main-chain carbonyl or amide groups between residues Val^{23} and Leu^{31} being involved in hydrogen bonding.

Flavivirus-encoded proteases also require the association of a cofactor for optimal catalytic activity (Falgout *et al.*, 1991; Bartenschlager *et al.*, 1994). However, this cofactor requirement is supplied for flavivirus proteases either in *cis* or *trans* by upstream sequences of the NS2B protein. Detailed biochemical studies and deletion analyses have mapped the region of NS2B

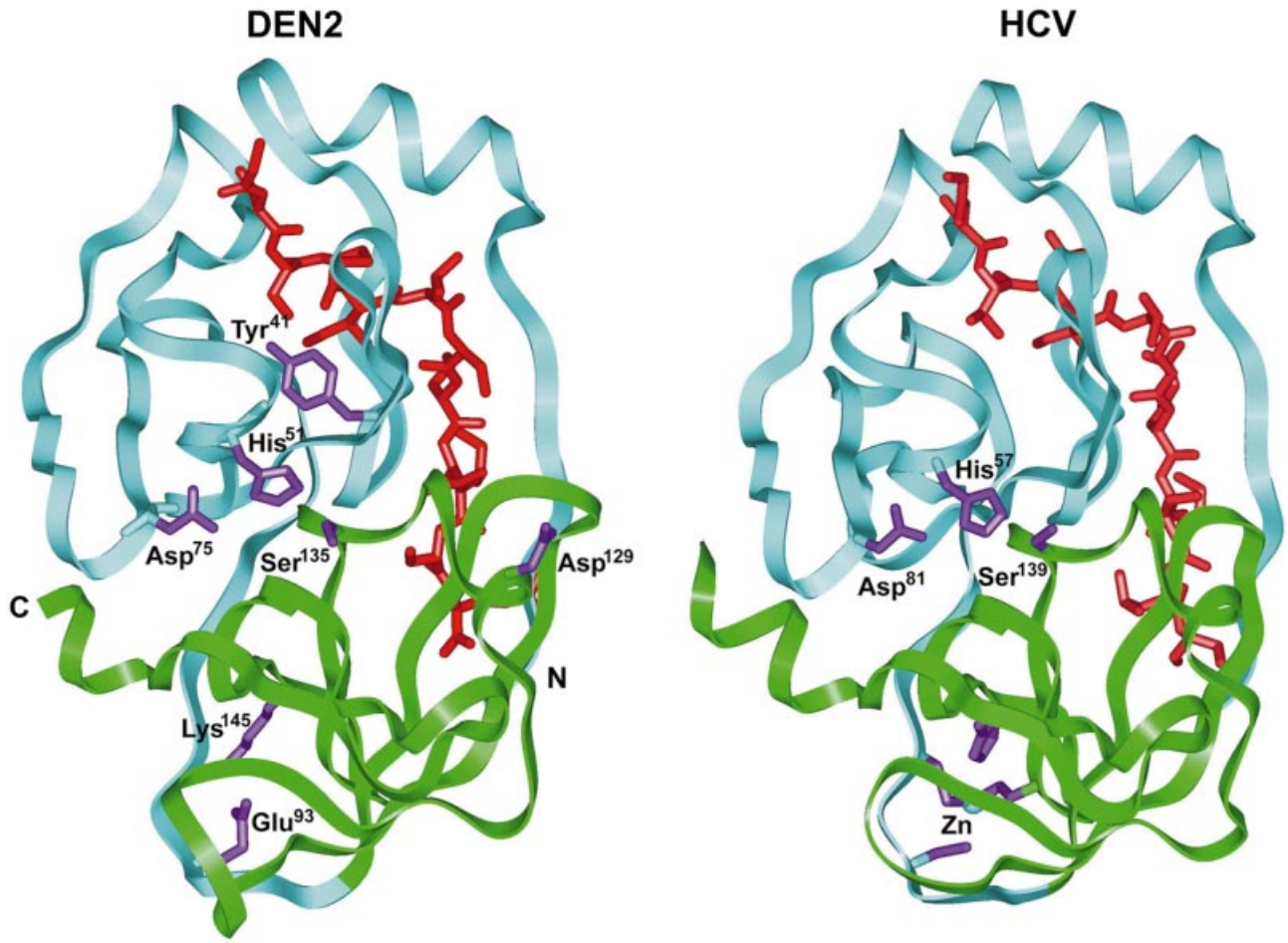


Fig. 3. Comparison of the backbone structures of the homology model of DEN2pro and the crystal structure of HCVpro complexed with their respective cofactors (the latter in red). Domain 1 is in blue and Domain 2 is in green (see text for details). The catalytic triad is identified in both structures, as are the residues involved in Zn-binding in HCVpro and the corresponding putative salt bridge in DEN2pro (purple).

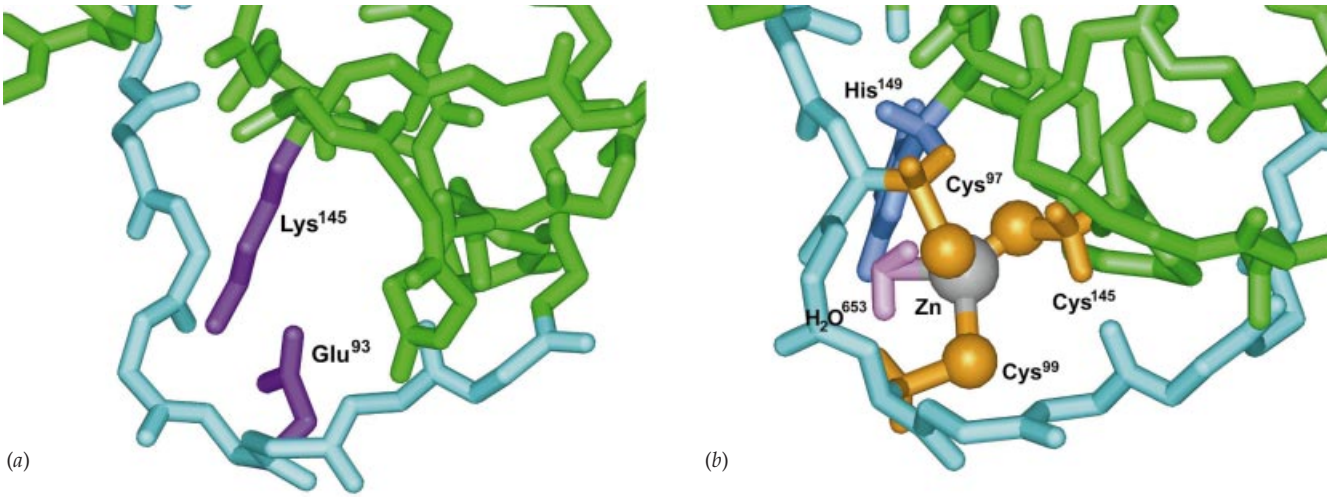
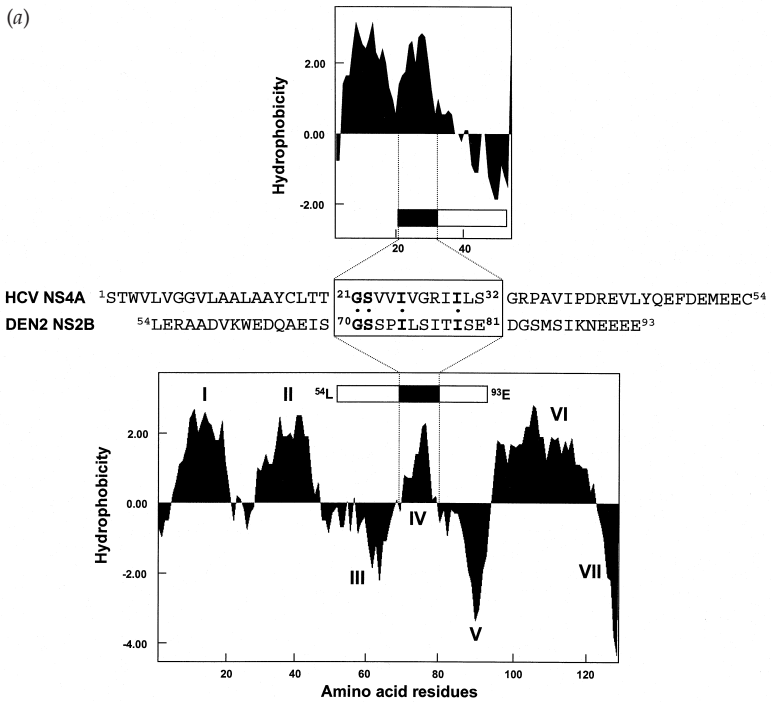


Fig. 4. Close-up view of (*a*) putative salt bridge in DEN2pro and (*b*) bonds between HCVpro and the complexed zinc ion. The backbone residues of domain 1 are in blue and those of domain 2 are in green.



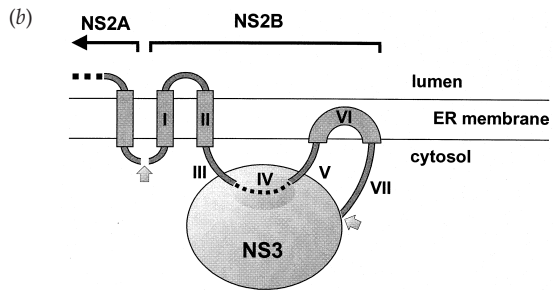


Fig. 5. Cofactor hydrophobicity plots and a hypothetical model of NS2B–NS3pro association with membranes. (*a*) Comparative hydrophobicity profiles [generated by the Kyte & Doolittle algorithm (Kyte & Doolittle, 1982) in the MacVector suite of programs] of the HCV cofactor NS4A (upper panel) and the DEN2 cofactor NS2B (lower panel). I–VII, regions of NS2B with either hydrophobic (I, II, IV and VI) or hydrophilic (III, V and VII) character. The complete amino acid sequence of HCV NS4A and the 40 amino acid cofactor domain of NS2B (schematically represented as a boxed region in the lower panel, with related regions similarly identified for NS4A in the upper panel) are also indicated. The central hydrophobic 12 amino acid stretch of NS2B modelled in this study is boxed, along with the region of the HCV NS4A cofactor known to interact with HCVpro in the crystal structure (identical residues are in bold). (*b*) Schematic representation of potential interactions between specific domains of the NS2B cofactor and either cellular membranes or NS3pro. Domains I–VII correspond to regions of hydrophobicity and hydrophilicity identified in (*a*). Shaded arrows identify sites of NS2B/NS3 cleavage.

required for NS3pro activity to a central, hydrophilic 40 amino acid domain (Lys^{54}–Glu^{93}; Fig. 5 *a*, lower panel) in an otherwise relatively hydrophobic protein (Falgout *et al.*, 1993; Clum *et al.*, 1997).

In searching for a shorter cofactor region of the dengue virus NS2B analogous to the minimum NS4A cofactor sequence of HCVpro, we focussed on a hydrophobic stretch of residues, Gly^{70}–Glu^{81}, located within the previously determined hydrophilic cofactor domain (Fig. 5 *a*, lower panel). A similarly located hydrophobic region is present in the NS2B protein of all flaviviruses. This region also shared some sequence similarity with the HCV NS4A protein (highlighted in bold in Fig. 5 *a*). We used this sequence to model the NS2B cofactor interaction with DEN2pro (Fig. 6 *a*). The model was constructed *in situ* as part of the complex, so docking of the two models was unnecessary. The interactions between DEN2pro and the NS2B cofactor are similar to those described for the interactions between HCVpro and its cofactor, with most of the interactions involving homologous residues in the two complexes (Figs 3 and 6 *a*). The putative NS2B cofactor is marginally more hydrophilic than its NS4A counterpart, which is reflected in some novel, mostly hydrogen-bond interactions. The model predicts that the phenolic hydroxyl of Tyr^{41} forms a hydrogen bond with the hydroxyl of Ser^{72} of the NS2B cofactor, while the methylene of Ser^{72} sits in a hydrophobic pocket made up of Leu^{58}, Tyr^{79} and Ile^{36}. Leu^{58} may play a central role in forming the hydrophobic pocket, and is conserved in HCV as Leu^{62}.

For HCVpro, Val^{105} and Val^{31} form part of another hydrophobic pocket, which interacts with Ile^{29} of the NS4A cofactor. The equivalent of Val^{105} in DEN2pro is Glu^{101}, and this may be hydrogen-bonded to the hydroxyl of Thr^{78} of the cofactor, while the methyl group of this residue could bind to a hydrophobic pocket consisting of Leu^{31} (the equivalent of Val^{31} of HCVpro) and Pro^{106}. Glu^{101} may also form a hydrogen bond with the hydroxyl of Ser^{80} of the cofactor. Other hydrogen-bond interactions include the main-chain carbonyl of Val^{7} with the hydroxyl of Ser^{76}. Ser^{76} also hydrogen bonds with the methylene of Tyr^{33}. Other potential hydrophobic NS2B cofactor–DEN2pro interactions are Pro^{73}–Met^{59}, Leu^{75}–Pro^{10}, Ile^{77}–Val^{7} and Ile^{79}–Val^{7} and –Trp^{5}.

Experiments examining the role of NS2B and NS4A cofactors indicate that in addition to activating proteolytic activity, they both serve to promote membrane association of the NS3 complex (Tanji *et al.*, 1995; Clum *et al.*, 1997). Indeed, cryoimmunoelectron microscopy studies have recently suggested that functional NS2B–NS3 proteolytic activity may in fact be compartmentalized to specific membranous structures (Westaway *et al.*, 1997). For HCV, the 20 N-terminal hydrophobic residues of NS4A (Fig. 5 *a*, upper panel) are predicted to form a membrane-spanning helix (Rost *et al.*, 1995) that could serve to anchor NS4A and the protease

complex to the cellular membrane. A similar mechanism may operate for NS2B, although it would be slightly more complex.

The interaction of both NS3pro and cellular membranes with NS2B may be described in terms of the hydrophobicity profile of NS2B (Fig. 5). There are seven major domains in NS2B that can be separated on the basis of their relative hydrophobicity (domains I to VII in Fig. 5*a*, *b*). As the NS2A/NS2B junction is cleaved by the viral protease located on the cytosolic side of the cellular membrane, this site must be similarly orientated. The first two hydrophobic domains (I and II) of NS2B are likely to function as a double anchor that would tether the protein in the membrane and also deliver the central cofactor domain (represented by the boxed region in the lower panel of Fig. 5*a*) to the cytosol for interaction with NS3pro. The core hydrophobic residues (domain IV) proposed in this study to interact with NS3pro are flanked by two hydrophilic stretches (domains III and V) that together comprise the recognized cofactor domain (Leu^{54}–Glu^{93}; Falgout *et al.*, 1993).

A direct comparison with NS4A reveals that domain V shares a highly charged, acidic character (rich in Glu residues) with downstream hydrophilic sequences in NS4A (open boxes in Fig. 5*a*). However, it is interesting to note the absence of a stretch of hydrophilic residues N-terminal of the core NS4A cofactor domain, as is found in NS2B (domain III). The importance of this upstream domain was clearly demonstrated in both cell-free translation and transient expression studies of deletion mutants of the YFV NS2B (Chambers *et al.*, 1993). In that study, the deletion of as few as three or five residues surrounding the N terminus of this hydrophilic domain abolished cofactor activity.

The specific function of these flanking sequences in NS2B is unknown. However, it is possible that they may contribute additional and essential points of contact between the protease complex and substrate. Alternatively, their function may be simply to present the central hydrophobic core for interaction with NS3pro. It should be noted that the majority of published studies that have analysed the minimal sequence of NS2B required for NS3pro activation have involved the expression of deletion mutants. It is possible that deletions within the flanking hydrophilic regions merely alter the tertiary folding of NS2B such that the hydrophobic domain identified in this study is not accessible for association with NS3pro. *In vitro* protease activation studies, similar to those carried out for NS4A (Lin *et al.*, 1995; Butkiewicz *et al.*, 1996), with small synthetic peptide cofactors and substrates have yet to be performed for the flavivirus NS2B. Our model would predict that by analogy with NS4A, NS2B peptides encompassing residues Gly^{70}–Gly^{83} may be sufficient to function as a peptide cofactor *in vitro*. The remaining domains identified in the hydrophobic profile of NS2B include a relatively long stretch of hydrophobic residues (domain VI), which is likely to function as an intramembrane domain similar to that of the caveolins (Parton *et al.*, 1996). The short hydrophilic stretch at the C terminus of NS2B (domain VII) would then serve to deliver the NS2B/NS3 junction to the activated protease for *cis* cleavage.

Another function of the NS2B cofactor may be to link and stabilize the two domains of the enzyme. Energy minimization of DEN2pro alone results in the two domains drifting apart, with the catalytic serine moving away from the other members of the catalytic triad. In this context, it is worth noting the differences in the crystal structures of HCVpro determined in the presence (Kim *et al.*, 1996; Yan *et al.*, 1998) and absence (Love *et al.*, 1996) of the NS4A peptide cofactor. In the latter structure, one member of the catalytic triad, Asp^{81}, is orientated away from the other two residues. A hypothetical model of the kinetics of NS4A complex formation with HCVpro has recently been proposed (Love *et al.*, 1998) and provides a structural explanation for the substantial tertiary rearrangements, particularly in the N terminus of HCVpro, induced during cofactor binding.

Substrate modelling

The substrate specificity of the flavivirus protease is fundamentally different from that of HCVpro. While HCVpro recognizes cysteine or threonine residues at the P1 position in the substrate, the flavivirus equivalent recognizes a pair of basic residues in the P1 and P2 positions. In both cases, the cleavage site is followed by a small, non-branched amino acid at P1′ (Rice *et al.*, 1986; Biedrzycka *et al.*, 1987; Speight *et al.*, 1988; Grakoui *et al.*, 1993). It is not surprising, therefore, that modelling of substrate interactions with DEN2pro (Fig. 6*b*) reveals differences from those proposed for HCVpro (Kim *et al.*, 1996; Love *et al.*, 1996). Nevertheless, both enzymes share structural features in common with other chymotrypsin-like proteases at the site of substrate interaction. The locations of the catalytic triad residues are highly conserved, as are the positions of the main-chain amides Gly^{133} and Ser^{135} (Gly^{137} and Ser^{139} in HCVpro; Gly^{193} and Ser^{195} in chymotrypsin) that form the oxyanion hole. The twisted β-strand, βE2, that parallels the substrate backbone and contributes to substrate binding (Edwards & Bernstein, 1994) is also well conserved between the proteases. The relative importance of this strand in substrate interactions is evidenced both by the almost complete conservation of residues amongst the flaviviruses in this region (Tyr^{150}–Val^{154}; Fig. 2) and by recent mutagenesis studies demonstrating loss of catalytic activity with all but the most conservative of amino acid substitutions (Valle & Falgout, 1998).

In common with HCVpro, DEN2pro has substantially shorter loop regions between βE1 and the catalytic Asp^{75} and between βB2 and βC2 than other members of the chymotrypsin family. These loop regions are involved in crucial contacts with substrate residues (Kim *et al.*, 1996) and their absence from the viral proteases results in the substrate being essentially ‘spread-eagled’ over the face of a relatively solvent-exposed binding domain (Fig. 6*b*, *c*). There are no substantial pockets for the

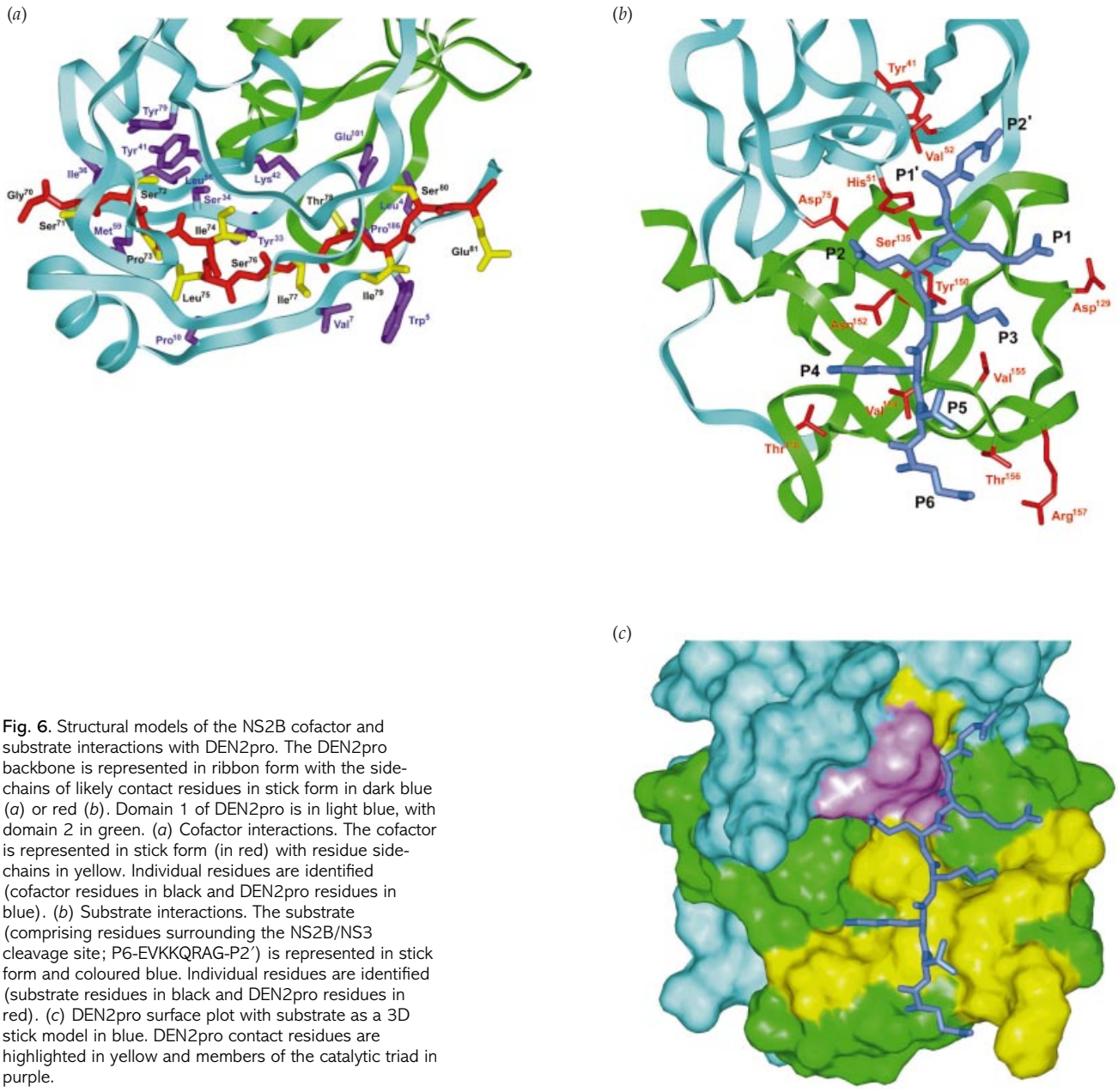


Fig. 6. Structural models of the NS2B cofactor and substrate interactions with DEN2pro. The DEN2pro backbone is represented in ribbon form with the side-chains of likely contact residues in stick form in dark blue (*a*) or red (*b*). Domain 1 of DEN2pro is in light blue, with domain 2 in green. (*a*) Cofactor interactions. The cofactor is represented in stick form (in red) with residue side-chains in yellow. Individual residues are identified (cofactor residues in black and DEN2pro residues in blue). (*b*) Substrate interactions. The substrate (comprising residues surrounding the NS2B/NS3 cleavage site; P6-EVKKQRAG-P2′) is represented in stick form and coloured blue. Individual residues are identified (substrate residues in black and DEN2pro residues in red). (*c*) DEN2pro surface plot with substrate as a 3D stick model in blue. DEN2pro contact residues are highlighted in yellow and members of the catalytic triad in purple.

amino acid side-chains of the substrate to bind to but rather, it appears to cling to the surface of the enzyme in an extended β-strand interaction (Fig. 6 *c*). Indeed, detailed analyses of hundreds of serine, cysteine, metallo- and aspartic proteases reveals that all such proteases recognize substrates and inhibitors in only the extended conformation (Reid & Fairlie, 1997; Fairlie *et al.*, 1998). This explains why folded elements of secondary structure are resistant to proteolytic degradation while denatured proteins or unfolded segments are degraded.

The substrate-cleavage site modelled in this study covers the junction between the DEN2 virus proteins NS2B and NS3 (EVKKQR ↓ AG). Structural comparisons of the P1-binding sites of members of the chymotrypsin-fold serine proteases suggest that substrate specificity is determined by a cluster of residues located near the catalytic serine (Perona & Craik, 1995). For HCVpro, the small hydrophobic specificity pocket for P1 is formed by the residues Leu^{135}, Phe^{154} and Ala^{157} (Love *et al.*, 1996). The equivalent residues in DEN2pro are Ser^{131}, Tyr^{150} and Gly^{153} (Fig. 2). However, in our model, Ser^{131} does not appear to play a role in substrate binding, a finding which is consistent with both a lack of conservation of this residue amongst the flaviviruses (Fig. 2) and mutagenesis studies which indicate that substitutions at this site do not affect proteolytic activity (Valle & Falgout, 1998). Tyr^{150} appears to lie beneath the backbone of the substrate as a sort of platform under the β-methylenes of the P1 and P3 side-

Table 2. Putative interactions between the dengue 2 virus NS2B/NS3 substrate cleavage site (P6-EVKKQRA-P1′) and the NS3 protease

Side-chain interactions are with main functional groups unless otherwise stated. H-bond, hydrogen bond.

Substrate residue		NS3pro residue	Interaction
P6	Glu	Arg^{157} or Thr^{156}(OH)	Salt bridge/H-bond
P5	Val	Val^{155}	Hydrophobic
		Thr^{156} (CH_3)	Hydrophobic
	Val CO	Val^{155} (NH)	H-bond*
P4	Lys-NH_2	Thr^{118} (OH)	H-bond
	Lys (βCH_2)	Val^{154}	Hydrophobic
P3	Lys	Ser^{131} or Asp^{129}	Salt bridge/H-bond
	Lys NH	Gly^{153} C=O	H-bond*
	Lys C=O	Gly^{153} NH	H-bond*
P2	Gln	Asp^{75}	H-bond
	Gln CO	Leu^{128}	Van der Waals
P1	Arg	Asp^{129}	Salt bridge/H-bond
P1′	Ala	Val^{52}	Hydrophobic
		Tyr^{41} (βCH_2)	Hydrophobic

* Backbone hydrogen bonds.

chains, with its phenolic hydroxyl possibly hydrogen bonding with one of the backbone components. This residue is highly conserved in the flaviviruses, with mutagenesis studies revealing that only a conservative Phe substitution is tolerated (Valle & Falgout, 1998). As with Ser^{131}, Gly^{153} does not appear to be involved in the definition of substrate specificity for DEN2pro. Nevertheless, the amino and carbonyl groups of Gly^{153} are within hydrogen-bonding distance of P3, suggesting the importance of this residue in overall substrate binding. Again, mutagenesis studies support this observation, with substitutions not being tolerated.

The substrate specificity of DEN2pro at the P1 position is similar to that of trypsin in recognizing basic lysine or arginine residues. This is defined for trypsin by an acidic Asp residue at position 189, six residues before the catalytic Ser^{195}, which forms a salt bridge with the lysine or arginine. Significantly, DEN2pro has an equivalent Asp at position 129, six residues before the catalytic Ser^{135}, that is totally conserved in all flavivirus NS3 sequences determined to date. Based on observations with the HCVpro crystal structure, Love *et al.* (1996) suggested that this Asp residue may be located at the bottom of the P1 pocket for the flavivirus NS3, as is the case for thrombin. Our model suggests, however, that Asp^{129} lies at the end of a P1-binding trough (Fig. 6*c*). The conformation of the α_b helix (Gly^{121}–Gly^{132}) in DEN2pro is different from that of the equivalent region of HCVpro, Arg^{119}–Pro^{131}, with adjustments in the orientation of Asp^{129} being required to bring it within hydrogen-bonding distance of P1. It should be noted however that substitutions in this residue are well tolerated. Valle & Falgout (1998) suggested that these results could be explained by the substituted amino acids still retaining contact with the substrate via a water molecule. Alternatively, given the specificity of flavivirus proteases for two basic residues at the P1 and P2 positions, the loss of one contact may not be sufficient to eliminate activity. Valle & Falgout (1998) speculate further that the second basic residue may be recognized by another Asp residue outside homology boxes 3 or 4. It is interesting to note in this context that the catalytic triad member Asp^{75} is indeed within hydrogen-bond distance of P2. The interaction of P2 with Asp^{75} would provide additional structural stability for the enzyme, given the link this bond forms between the two protein domains (Fig. 6*b*). During energy minimization of the enzyme–cofactor complex model, it was found that in the absence of substrate the two domains of the enzyme, and hence the catalytic triad residues, moved apart. A precedent for the catalytic Asp contacting an additional residue is provided by HCVpro itself, in which the catalytic Asp^{81} forms an ion pair with Arg^{115} in the βE2 strand (Kim *et al.*, 1996).

Additional interactions between DEN2pro and substrate that are predicted by this model are summarized in Table 2. The DEN2pro residues Tyr^{41}, Val^{52}, Asp^{75}, Thr^{118}, Asp^{129}, Gly^{153}, Val^{155}, Thr^{156} and Arg^{157} are all likely to contribute to binding of substrate residues between P2′ and P6 (Fig. 6*b*). For the cleavage site modelled in this study (the NS2B/NS3 junction), the P4 Lys hydrogen-bonds with the hydroxyl of Thr^{118}. Other cleavage sites in the DEN2 virus polyprotein recognized by the NS2B–NS3 protease have an Ala (NS3/NS4A) or Ser (NS2A/NS2B and NS4A and NS5) in the P4 position. Val^{154} is well placed to bind to these side-chains with β-carbons, since it lies close to, and appears to bind to, the β-methylene of Lys P4 in the model.

Concluding remarks

This study describes a structural model for DEN2pro based upon homology to the crystal structure of HCVpro. The principal aim of this study was to gain some insight into the structural basis of substrate processing by DEN2pro that could aid inhibitor design. The model predicts novel features for both substrate and cofactor binding with the protease. During the development of the model, energy minimization analyses clearly demonstrated a role for both substrate and cofactor in the stabilization of the structure of NS3pro. For example, interactions between the substrate and enzyme residues spanning the binding cleft separating the two domains of NS3pro suggest that the substrate may act as a clasp to maintain association of the two domains. The model also suggests a reason for the difference between the dengue virus and HCV proteases in their requirement for metal ions. The metal-independent protease DEN2pro may have a structure-stabilizing salt bridge in place of the zinc-binding motif present in the zinc-dependent protease HCVpro. Our model also predicted that the core component of the NS2B cofactor that directly associates with NS3 may be as short as 12 residues, rather than the 40 amino acid cofactor domain which has been previously identified, and provides a structural basis for understanding the association of DEN2pro with its NS2B cofactor. By analogy with the NS4A cofactor for HCVpro, NS2B residues Gly^{70}–Gly^{83} may be the minimum requirement for a peptide cofactor in *in vitro* cleavage assays. Experimental studies are now in progress to test the structural conclusions and substrate/cofactor requirements deduced from this study.

This work was supported in part by the Royal Children's Hospital Foundation (to P.R.Y.). This is publication #92 from the Sir Albert Sakzewski Virus Research Centre.

References

Arias, C. F., Preugschat, F. & Strauss, J. H. (1993). Dengue 2 virus NS2B and NS3 form a stable complex that can cleave NS3 within the helicase domain. *Virology* **193**, 888–899.

Bartenschlager, R., Ahlborn-Laake, L., Mous, J. & Jacobsen, H. (1994). Kinetic and structural analyses of hepatitis C virus polyprotein processing. *Journal of Virology* **68**, 5045–5055.

Bazan, J. F. & Fletterick, R. J. (1989). Detection of a trypsin-like serine protease domain in flaviviruses and pestiviruses. *Virology* **171**, 637–639.

Bazan, J. F. & Fletterick, R. J. (1990). Structural and catalytic models of trypsin-like viral proteases. *Seminars in Virology* **1**, 311–322.

Biedrzycka, A., Cauchi, M. R., Bartholomeusz, A., Gorman, J. J. & Wright, P. J. (1987). Characterization of protease cleavage sites involved in the formation of the envelope glycoprotein and three non-structural proteins of dengue virus type 2, New Guinea C strain. *Journal of General Virology* **68**, 1317–1326.

Butkiewicz, N. J., Wendel, M., Zhang, R., Jubin, R., Pichardo, J., Smith, E. B., Hart, A. M., Ingram, R., Durkin, J., Mui, P. W., Murray, M. G., Ramanathan, L. & Dasmahapatra, B. (1996). Enhancement of hepatitis C virus NS3 proteinase activity by association with NS4A-specific synthetic peptides: identification of sequence and critical residues of NS4A for the cofactor activity. *Virology* **225**, 328–338.

Castle, E., Leidner, U., Nowak, T., Wengler, G. & Wengler, G. (1986). Primary structure of the West Nile flavivirus genome region coding for all nonstructural proteins. *Virology* **149**, 10–26.

Chambers, T. J., Hahn, C. S., Galler, R. & Rice, C. M. (1990*a*). Flavivirus genome organization, expression, and replication. *Annual Review of Microbiology* **44**, 649–688.

Chambers, T. J., Weir, R. C., Grakoui, A., McCourt, D. W., Bazan, J. F., Fletterick, R. J. & Rice, C. M. (1990*b*). Evidence that the N-terminal domain of nonstructural protein NS3 from yellow fever virus is a serine protease responsible for site-specific cleavages in the viral polyprotein. *Proceedings of the National Academy of Sciences, USA* **87**, 8898–8902.

Chambers, T. J., Nestorowicz, A., Amberg, S. M. & Rice, C. M. (1993). Mutagenesis of the yellow fever virus NS2B protein: effects on proteolytic processing, NS2B–NS3 complex formation, and viral replication. *Journal of Virology* **67**, 6797–6807.

Choi, H.-K., Tong, L., Minor, W., Dumas, P., Boege, U., Rossmann, M. G. & Wengler, G. (1991). Structure of Sindbis virus core protein reveals a chymotrypsin-like serine proteinase and the organization of the virion. *Nature* **354**, 37–43.

Choi, H.-K., Lu, G., Lee, S., Wengler, G. & Rossmann, M. G. (1997). Structure of Semliki Forest virus core protein. *Proteins* **27**, 345–359.

Clum, S., Ebner, K. E. & Padmanabhan, R. (1997). Cotranslational membrane insertion of the serine proteinase precursor NS2B–NS3(Pro) of dengue virus type 2 is required for efficient in vitro processing and is mediated through the hydrophobic regions of NS2B. *Journal of Biological Chemistry* **272**, 30715–30723.

De Francesco, R., Urbani, A., Nardi, M. C., Tomei, L., Steinkuhler, C. & Tramontano, A. (1996). A zinc binding site in viral serine proteinases. *Biochemistry* **35**, 13282–13287.

Edwards, P. D. & Bernstein, P. R. (1994). Synthetic inhibitors of elastase. *Medicinal Research Reviews* **14**, 127–194.

Fairlie, D. P., West, M. L. & Wong, A. K. (1998). Towards protein surface mimetics. *Current Medicinal Chemistry* **5**, 29–62.

Falgout, B. & Markoff, L. (1995). Evidence that flavivirus NS1–NS2A cleavage is mediated by a membrane-bound host protease in the endoplasmic reticulum. *Journal of Virology* **69**, 7232–7243.

Falgout, B., Pethel, M., Zhang, Y.-M. & Lai, C.-J. (1991). Both nonstructural proteins NS2B and NS3 are required for the proteolytic processing of dengue virus nonstructural proteins. *Journal of Virology* **65**, 2467–2475.

Falgout, B., Miller, R. H. & Lai, C.-J. (1993). Deletion analysis of dengue virus type 4 nonstructural protein NS2B: identification of a domain required for NS2B–NS3 protease activity. *Journal of Virology* **67**, 2034–2042.

Fu, J., Tan, B.-H., Yap, E.-H., Chan, Y.-C. & Tan, Y. H. (1992). Full-length cDNA sequence of dengue type 1 virus (Singapore strain S275/90). *Virology* **188**, 953–958.

Gorbalenya, A. E., Donchenko, A. P., Koonin, E. V. & Blinov, V. M. (1989). N-terminal domains of putative helicases of flavi- and pestiviruses may be serine proteases. *Nucleic Acids Research* **17**, 3889–3897.

Grakoui, A., McCourt, D. W., Wychowski, C., Feinstone, S. M. & Rice, C. M. (1993). Characterization of the hepatitis C virus-encoded serine proteinase: determination of proteinase-dependent polyprotein cleavage sites. *Journal of Virology* **67**, 2832–2843.

Kim, J. L., Morgenstern, K. A., Lin, C., Fox, T., Dwyer, M. D., Landro, J. A., Chambers, S. P., Markland, W., Lepre, C. A., O'Malley, E. T.,

Harbeson, S. L., Rice, C. M., Murcko, M. A., Caron, P. R. & Thomson, J. A. (1996). Crystal structure of the hepatitis C virus NS3 protease domain complexed with a synthetic NS4A cofactor peptide. *Cell* **87**, 343–355.

Kyte, J. & Doolittle, R. F. (1982). A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology* **157**, 105–132.

Lin, C., Amberg, S. M., Chambers, T. J. & Rice, C. M. (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. *Journal of Virology* **67**, 2327–2335.

Lin, C., Thomson, J. A. & Rice, C. M. (1995). A central region in the hepatitis C virus NS4A protein allows formation of an active NS3–NS4A serine proteinase complex in vivo and in vitro. *Journal of Virology* **69**, 4373–4380.

Lobigs, M. (1993). Flavivirus premembrane protein cleavage and spike heterodimer secretion require the function of the viral proteinase NS3. *Proceedings of the National Academy of Sciences, USA* **90**, 6218–6222.

Love, R. A., Parge, H. E., Wickersham, J. A., Hostomsky, Z., Habuka, N., Moomaw, E. W., Adachi, T. & Hostomska, Z. (1996). The crystal structure of hepatitis C virus NS3 proteinase reveals a trypsin-like fold and a structural zinc binding site. *Cell* **87**, 331–342.

Love, R. A., Parge, H. E., Wickersham, J. A., Hostomsky, Z., Habuka, N., Moomaw, E. W., Adachi, T., Margosiak, S., Dagostino, E. & Hostomska, Z. (1998). The conformation of hepatitis C virus NS3 proteinase with and without NS4A: a structural basis for the activation of the enzyme by its cofactor. *Clinical and Diagnostic Virology* **10**, 151–156.

Lüthy, R., Bowie, J. U. & Eisenberg, D. (1992). Assessment of protein models with three-dimensional profiles. *Nature* **356**, 83–85.

Mackow, E., Makino, Y., Zhao, B. T., Zhang, Y.-M., Markoff, L., Buckler-White, A., Guiler, M., Chanock, R. & Lai, C.-J. (1987). The nucleotide sequence of dengue type 4 virus: analysis of genes coding for nonstructural proteins. *Virology* **159**, 217–228.

Nowak, T., Farber, P. M., Wengler, G. & Wengler, G. (1989). Analyses of the terminal sequences of West Nile virus structural proteins and of the *in vitro* translation of these proteins allow the proposal of a complete scheme of the proteolytic cleavages involved in their synthesis. *Virology* **169**, 365–376.

Osatomi, K. & Sumiyoshi, H. (1990). Complete nucleotide sequence of dengue type 3 virus genome RNA. *Virology* **176**, 643–647.

Parton, R. G. (1996). Caveolae and caveolins. *Current Opinion in Cell Biology* **8**, 542–548.

Perona, J. J. & Craik, C. S. (1995). Structural basis of substrate specificity in the serine proteases. *Protein Science* **4**, 337–360.

Preugschat, F., Yao, C.-W. & Strauss, J. H. (1990). In vitro processing of dengue virus type 2 nonstructural proteins NS2A, NS2B, and NS3. *Journal of Virology* **64**, 4364–4374.

Preugschat, F., Lenches, E. M. & Strauss, J. H. (1991). Flavivirus enzyme–substrate interactions studied with chimeric proteinases: identification of an intragenic locus important for substrate recognition. *Journal of Virology* **65**, 4749–4758.

Reid, R. C. & Fairlie, D. P. (1997). Mimicking extended conformations of protease substrates: designing cyclic peptidomimetics to inhibit HIV-1 protease. In *Advances in Amino Acid Mimetics and Peptidomimetics*, pp. 77–107. Edited by A. Abell. London: JAI Press.

Rice, C. M., Lenches, E. M., Eddy, S. R., Shin, S. J., Sheets, R. L. & Strauss, J. H. (1985). Nucleotide sequence of yellow fever virus: implications for flavivirus gene expression and evolution. *Science* **229**, 726–733.

Rice, C. M., Aebersold, R., Teplow, D. B., Pata, J., Bell, J. R., Vorndam, A. V., Trent, D. W., Brandriss, M. W., Schlesinger, J. J. & Strauss, J. H. (1986). Partial N-terminal amino acid sequences of three nonstructural proteins of two flaviviruses. *Virology* **151**, 1–9.

Rost, B., Casadio, R., Fariselli, P. & Sander, C. (1995). Transmembrane helices predicted at 95% accuracy. *Protein Science* **4**, 521–533.

Ryan, M. D., Monaghan, S. & Flint, M. (1998). Virus-encoded proteinases of the *Flaviviridae*. *Journal of General Virology* **79**, 947–959.

Speight, G., Coia, G., Parker, M. D. & Westaway, E. G. (1988). Gene mapping and positive identification of the non-structural proteins NS2A, NS2B, NS3, NS4B and NS5 of the flavivirus Kunjin and their cleavage sites. *Journal of General Virology* **69**, 23–34.

Stadler, K., Allison, S. L, Schalich, J. & Heinz, F. X. (1997). Proteolytic activation of tick-borne encephalitis virus by furin. *Journal of Virology* **71**, 8475–8481.

Sumiyoshi, H., Mori, C., Fuke, I., Morita, K., Kuhara, S., Kondou, J., Kikuchi, Y., Nagamatu, H. & Igarishi, A. (1987). Complete nucleotide sequence of the Japanese encephalitis virus genome RNA. *Virology* **161**, 497–510.

Tanji, Y., Hijikata, M., Satoh, S., Kaneko, T. & Shimotohno, K. (1995). Hepatitis C virus-encoded nonstructural protein NS4A has versatile functions in viral protein processing. *Journal of Virology* **69**, 1575–1581.

Teo, K. F. & Wright, P. J. (1997). Internal proteolysis of the NS3 protein specified by dengue virus 2. *Journal of General Virology* **78**, 337–341.

Valle, R. P. C. & Falgout, B. (1998). Mutagenesis of the NS3 protease of dengue virus type 2. *Journal of Virology* **72**, 624–632.

Wengler, G., Czaya, G., Färber, P. M. & Hegemann, J. H. (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. *Journal of General Virology* **72**, 851–858.

Westaway, E. G., Mackenzie, J. M., Kenney, M. T., Jones, M. K. & Khromykh, A. A. (1997). Ultrastructure of Kunjin virus-infected cells: colocalization of NS1 and NS3 with double-stranded RNA, and of NS2B with NS3, in virus-induced membrane structures. *Journal of Virology* **71**, 6650–6661.

World Health Organization (1996). *Dengue and Dengue Haemorrhagic Fever. Report N117.* Geneva: World Health Organization.

Yan, Y., Li, Y., Munshi, S., Sardana, V., Cole, J. L., Sardana, M., Steinkuehler, C., Tomei, L., De Francesco, R., Kuo, L. C. & 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 Science* **7**, 837–847.

Received 22 October 1998; Accepted 8 January 1999