Journal of General Virology (1998), **79**, 947–959. Printed in Great Britain

REVIEW ARTICLE

Virus-encoded proteinases of the *Flaviviridae*

Martin D. Ryan,[1] Susan Monaghan[1] and Mike Flint[2]

[1] School of Biomedical Sciences, University of St. Andrews, Irvine Building, North Street, St Andrews KY16 9AL, UK
[2] School of Animal and Microbial Sciences, University of Reading, Whiteknights, PO Box 228, Reading RG6 6AJ, UK

Introduction

In our review of picornavirus proteinases (Ryan & Flint, 1997) we referred to these viruses as 'the exemplars' of the polyprotein strategy. This was meant in the sense that all of the virus proteins are encoded in a single, long, open reading frame (ORF) and all of the processing events (as are currently understood) are mediated by virus proteinases, or by virus-specific proteins if not proteinases sensu stricto. Viruses within the *Flaviviridae* (the flavi-, pesti- and hepatitis C viruses) also encode all of their proteins in a single, long ORF (ranging between ~3400 to ~4000 codons) with a polyprotein architecture similar to that of the picornaviruses – the structural proteins in the N-terminal portion of the polyprotein whilst the replicative (non-structural or NS) proteins constitute the remainder (Fig. 1; reviewed by Chambers *et al.*, 1990*a*). The *Flaviviridae* are enveloped viruses that use host-cell proteinases (signalases) to process at multiple sites both in the structural protein precursor and at some sites in non-structural protein precursors whilst the remainder of the cleavages are mediated by a virus-encoded proteinase (NS3 protein; Fig. 1). Unlike the picornaviruses, therefore, polyproteins of viruses within the *Flaviviridae* are processed by a combination of host and virus proteinases (reviewed by Rice & Strauss, 1990; Dougherty & Semler, 1993).

The classical test for polyprotein processing in *cis* (intramolecular) or in *trans* (intermolecular) is to observe the effect of dilution on the polyprotein cleavages which occur during translation *in vitro*. Data from these types of experiment on various members of the *Flaviviridae* suggested that polyprotein processing was a highly regulated process; (i) processing occurred by a combination of proteolytic events both in *cis* and in *trans*, (ii) the proteolytic activity of the N-terminal domain of NS3 could be modified by the presence of other virus proteins and (iii) processing occurred in an ordered fashion (Preugschat *et al.*, 1990, 1991; Chambers *et al.*, 1990*b*; Falgout *et al.*, 1991; Preugschat & Strauss, 1991). Furthermore, the order of cleavages could be altered by site-directed mutagenesis of the substrate binding pocket of the proteinase (Preugschat *et al.*, 1991). The controlled biogenesis of proteins is achieved, therefore, either by the polyprotein folding to 'oblige' certain enzyme–substrate interactions (cleavages in *cis*), by normal enzyme–substrate binding considerations (rates of cleavage of different sites in *trans*) or by biochemical regulation of the enzymes' activity (cofactors).

Descriptions of the active sites of the major classes of proteinases were given previously (Ryan & Flint, 1997, and references therein). One purpose of these reviews is to provide sequence alignments to enable extrapolation of the structural data determined for one proteinase to other similar enzymes within the group. Due to the large number of sequences available for the NS3 proteinases, however, it is not practicable to present alignments including all sequences. Full sequence alignments are available by e-mail (martin.ryan@st-and.ac.uk) or by anonymous FTP from ftp.st-and.ac.uk/info/ftp/-pub/mdr1 (together with updated picornavirus super-group proteinase alignments).

Author for correspondence: Martin Ryan.
Fax +44 1334 463400. e-mail martin.ryan@st-and.ac.uk

NS3 Proteinase

(i) Identification of the NS3 proteolytic domain

Analysis of sequence alignments predicted the existence of a trypsin-like serine proteinase domain within the N-terminal region of the flavi- and pestivirus NS3 protein. The proposed proteolytic domain was of some 180 aa with a catalytic triad, conserved in all sequences, of His-Asp-Ser (Bazan & Fletterick, 1989, 1990; Gorbalenya *et al.*, 1989*a*, *b*). A serine residue was predicted as the active site nucleophile in accordance with previous inhibitor studies showing inhibition of polyprotein processing by N^4-tosylphenylalanine chloromethyl ketone (TPCK; Cleaves, 1985). The alignment shown in Fig. 2 uses the N termini of NS3 experimentally determined for the flavi- and hepatitis C viruses. The N termini of the pestivirus sequences were chosen to maximize the alignment, the C termini of all sequences being arbitrary. The $NS3^{pro}$ phylogenetic tree shown in Fig. 3 shows the same pattern of relatedness as was found for a region of the helicase domain of NS3 (Ohba *et al.*, 1996).

The predictions of biochemical properties based on alignments were tested by analysis of the endogenous proteolytic properties of proteins derived from subgenomic cDNA clones encoding NS2A, NS2B and NS3 sequences derived from yellow fever (Chambers *et al.*, 1990*b*), dengue type 2

0001-5134

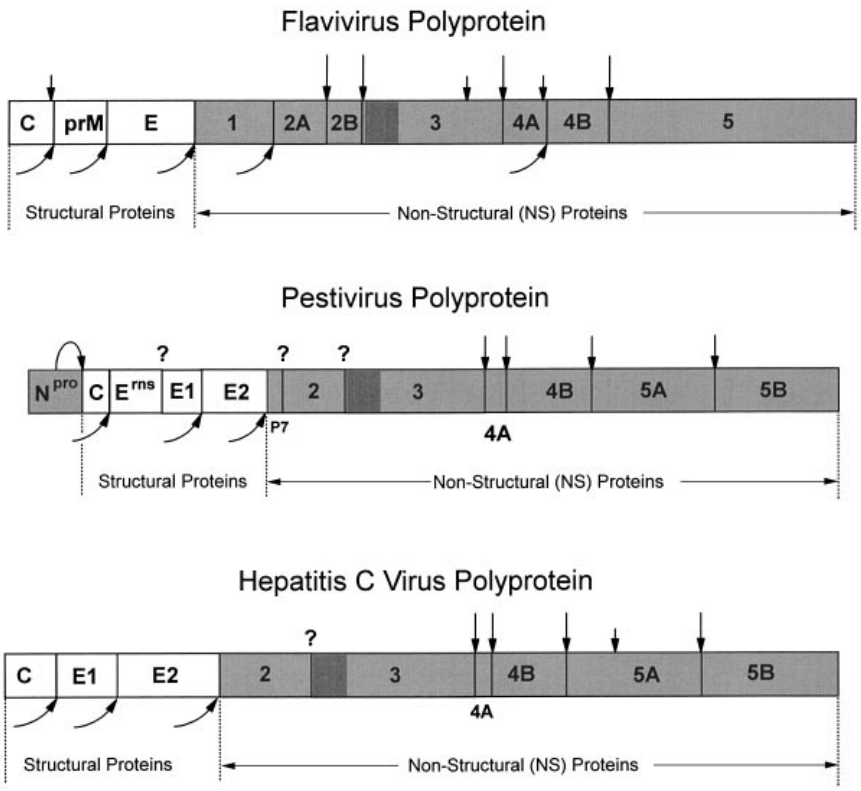

Fig. 1. Polyproteins of the *Flaviviridae*. Polyproteins are shown (boxed areas). Non-structural regions are shaded and the proteolytic domain of NS3 shown (darker shading). Host-cell proteinase-mediated cleavages are shown (curved arrows) together with the $NS3^{pro}$-mediated cleavages (vertical arrows). In the case of the flavi- and hepatitis C viruses the $NS3^{pro}$-mediated cleavages occurring *within* proteins are differentiated by the shorter vertical arrows. Cleavages mediated by uncharacterized proteinases are indicated by '?'.

(Preugschat *et al.*, 1990) or West Nile viruses (Wengler *et al.*, 1991). These studies showed that the proteinase domain was, indeed, located within the N-terminal ~ 180 residues of NS3 (referred to here as $NS3^{pro}$) and was responsible for cleavage at the NS2A/2B and NS2B/NS3 sites in an apparent intramolecular fashion. Site-directed mutagenesis experiments confirmed the importance of those residues predicted to form the catalytic triad: His^{51}, Asp^{75} and Ser^{135} (Japanese encephalitis virus numbering scheme in Fig. 2; Chambers *et al.*, 1990*b*; Wengler *et al.*, 1991; Pugachev *et al.*, 1993).

Analyses of hepatitis C and GB virus NS3 sequences showed a close relationship with flavi- and pestivirus NS3 sequences (Figs 2 and 3). Those residues identified previously as forming the catalytic triad of the flavivirus NS3 proteinase domain were conserved in alignments with hepatitis C and GB virus NS3 sequences – the strong inference being that the N-terminal domain of NS3 of these viruses would also possess proteolytic activity (Miller & Purcell, 1990; Choo *et al.*, 1991; see Fig. 2). Experimental evidence proved this to be the case for both hepatitis C and GB viruses (Bartenschlager *et al.*, 1993; Eckart *et al.*, 1993; Grakoui *et al.*, 1993*a*, *b*; Hijikata *et al.*, 1993*a*; Tomei *et al.*, 1993; Manabe *et al.*, 1994; Han *et al.*, 1995; Scarselli *et al.*, 1997).

The remainder of the NS3 protein contains characteristic NTPase and helicase motifs (Koonin & Dolja, 1993) and both biochemical activities have been demonstrated (Wengler & Wengler, 1991, 1993; Suzich *et al.*, 1993; D'Souza *et al.*, 1995; Warrerner & Collet, 1995; Kim *et al.*, 1997).

(ii) Activation/stimulation of $NS3^{pro}$

Following the identification of the NS3 proteolytic activity, polyprotein processing studies showed that flavivirus $NS3^{pro}$ required NS2B protein, which could be supplied either in *cis* or in *trans*, for activity (Chambers *et al.*, 1991; Falgout *et al.*, 1991, 1993; Cahour *et al.*, 1992; Arias *et al.*, 1993). The region of the NS2B protein required for $NS3^{pro}$ activity was mapped to a 40 aa tract located 77–37 aa upstream of the NS3 N terminus (Falgout *et al.*, 1993). The NS2B protein shows low conservation amongst the flaviviruses and only 3 aa are completely conserved in this hydrophilic 40 aa region of NS2B. The association of NS2B and NS3, demonstrated by co-immunoprecipitation experiments, is mediated by this hydrophilic region (pers. comms cited in Falgout *et al.*, 1993).

An early indication that pestiviruses were not directly analogous to flaviviruses in this respect was provided by an analysis of bovine viral diarrhoea virus (BVDV) polyprotein processing (Wiskerchen & Collett, 1991). Processing at the NS5A/5B site (site '10') showed a requirement for another 'co-factor' protein (NS4A/p10). Similarly, later studies on the processing of hepatitis C virus polyprotein showed an intriguing difference with that of the flaviviruses. Rather than a requirement of (upstream) NS2B sequences for proteolytic activity, the hepatitis C virus $NS3^{pro}$ activity was enhanced by (downstream) NS4A sequences (Bartenschlager *et al.*, 1994, 1995; Failla *et al.*, 1994, 1995; Hahm *et al.*, 1995; Lin & Rice, 1995; Lin *et al.*, 1995; Overton *et al.*, 1995; Satoh *et al.*, 1995;

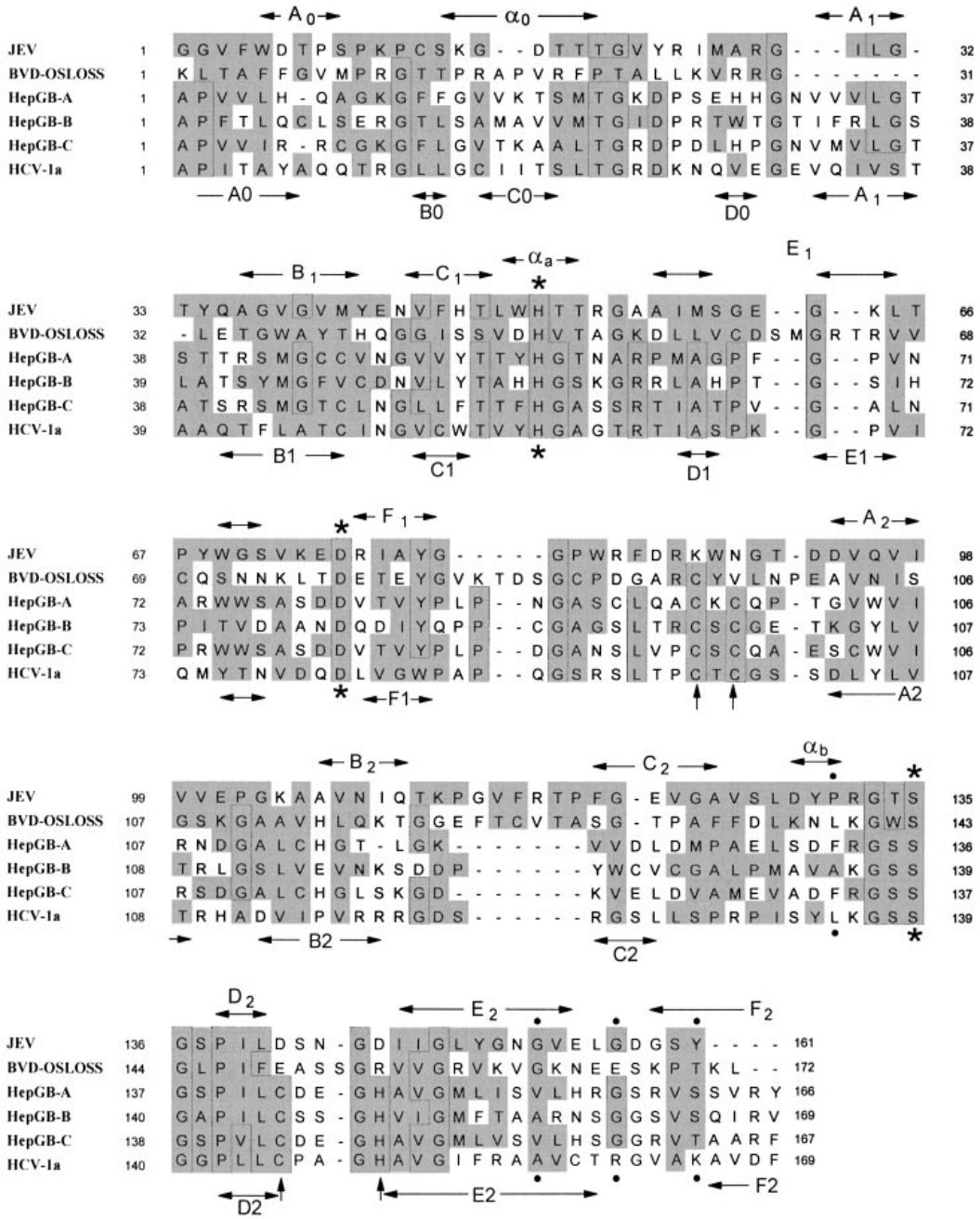

Fig. 2. Sequence alignment of flavi-, pesti-, hepatitis C and G virus NS3 proteolytic domains. Sequences of NS3 proteolytic domains listed below were aligned using CLUSTALW (Higgins *et al*., 1991). Active site residues are indicated by asterisks. Hepatitis C virus NS3 residues involved in binding zinc are indicated by arrows and residues involved in substrate binding are indicated by filled circles. Only a few NS3 sequences are shown here, a representative sequence being selected from the alignment for each major group of sequences (see Fig. 3). The N termini of the sequences are authentic cleavage sites (flavi- and hepatitis C viruses) or arbitrary positions determined by alignment (pesti-, hepatitis G viruses). All C termini are arbitrary, but reflect an appropriate region where the alignment is maintained across all the groups of viruses before the similarity drops too low. Sequences used in the analyses: dengue virus (type 1 – Fu *et al*., 1992; type 2 – Hahn *et al*., 1988; type 3 – Osatomi & Sumiyoshi, 1990; type 4 – Mackow *et al*., 1987); Japanese encephalitis virus (Sumiyoshi *et al*., 1987); Kunjin virus (Coia *et al*., 1988); Murray Valley encephalitis virus (MVEV – Dalgarno *et al*., 1986); tick-borne encephalitis virus (TBEV – Pletnev *et al*., 1990); West Nile virus (WNV Castle *et al*., 1986); yellow fever virus (YVF – Rice *et al*., 1985); hepatitis G virus (strain GB-A, B – Simons *et al*., 1995; C – Leary *et al*., 1995); hepatitis C virus (HCV strain 1a – Choo *et al*., 1991; 1b – Kato *et al*., 1990; 1c – Okamoto *et al*., 1994; 2a – Okamoto *et al*., 1991; 2b – Okamoto *et al*., 1992; 2c – Nakao *et al*., 1996; 3a – Sakamoto *et al*., 1994; 3b – Chayama *et al*., 1994; 4a – Chamberlain *et al*., 1997*a*; 5a – Chamberlain *et al*., 1997*b*; 6a – Adams *et al*., 1997); bovine viral diarrhoea virus (BVDV strain Osloss – Renard *et al*., 1987; De Moerlooze *et al*., 1993; NADL – Collet *et al*., 1988*b*; SD1 – Deng & Brock, 1992; NCP7 – Meyers *et al*., 1996; ILLC – Roath & Berry, unpublished; Il – Ridpath & Bolin, 1995); classical swine fever virus (CSFV strain Alfort – Meyers *et al*., 1989; Brescia – Moorman *et al*., 1990; C – Moorman *et al*., 1996); and border disease virus (BVD – Ridpath *et al*., unpublished).

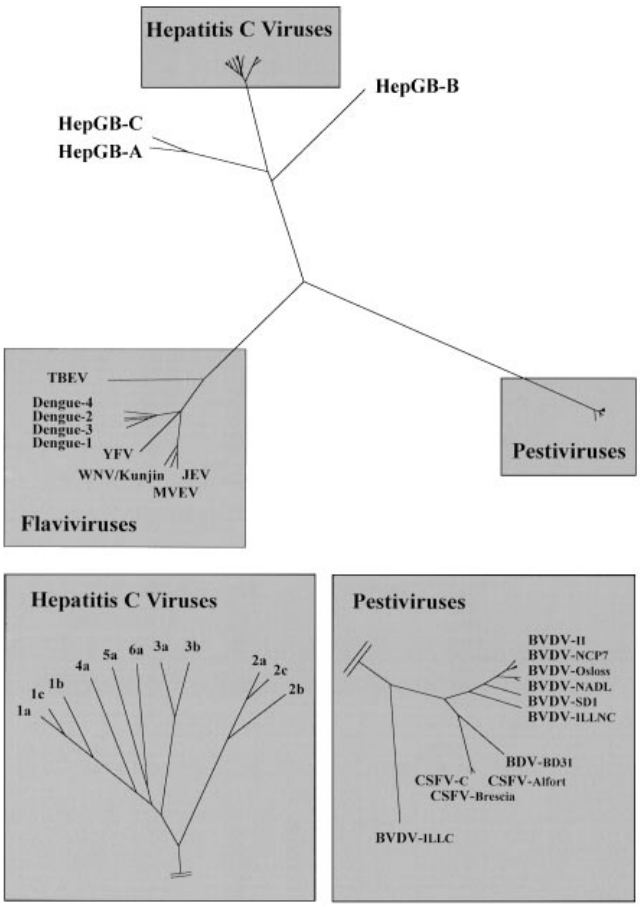

Fig. 3. Sequence similarities between NS3 proteolytic domains of the *Flaviviridae*. The sequence alignment described in the legend to Fig. 2 was analysed using PROTDIST and NEIGHBOUR to determine relationships. The results of the alignment of all sequences used were plotted using DRAWTREE (Felsenstein, 1991). Branch lengths are proportional to relatedness.

Tanji *et al.*, 1995; Koch *et al.*, 1996; Shimizu *et al.*, 1996; Tomei *et al.*, 1996). This cofactor activity has been mapped to a short oligopeptide sequence within the centre of NS4A (Butkiewicz *et al.*, 1996).

The use of assay systems *in vitro* (see below) has allowed the characterization of this NS4A stimulatory effect. The presence of an NS4A peptide stimulated NS3^{pro} activity by up to 100-fold – increasing the k_{cat} rather than the K_m of the enzyme (Shimizu *et al.*, 1996; Steinkuhler *et al.*, 1996*b*; Landro *et al.*, 1997). In the latter study, Landro *et al.* observed that the pH dependence of NS3^{pro} was not affected by the NS4A peptide – suggesting that NS4A does not alter the pK_a values of catalytic residues. Steady-state kinetic measurements indicated that the binding of the peptides was ordered during the catalytic cycle: activating NS4A peptide binds first, then substrate. Recent work has strongly implicated pestivirus NS4A as the region of the polyprotein responsible for activation of NS3^{pro} – similar to that observed in the case of hepatitis C viruses (Xu *et al.*, 1997).

The adenovirus 23K proteinase is also stimulated by a peptide cofactor – in this case an 11 aa peptide derived from the C terminus of protein pVI (Webster *et al.*, 1993). The k_{cat} of the enzyme is increased some 350-fold when incubated with an equimolar concentration of cofactor peptide (Mangel *et al.*, 1996). The crystal structure of the adenovirus 23K proteinase shows, however, that the peptide cofactor is bound at the surface, some distance from the active site (Ding *et al.*, 1996) – quite unlike the hepatitis C virus NS3^{pro}/NS4A structure (described below).

(iii) Analysis of NS3^{pro} activity

Analyses of the proteolytic activity of NS3^{pro} were simplified by the (relative) ease of the expression and purification of active enzyme using a wide range of heterologous systems: *Escherichia coli* (Tomei *et al.*, 1993; Komoda *et al.*, 1994; D'Souza *et al.*, 1995; Kakiuchi *et al.*, 1995; Bianchi *et al.*, 1996; Kim *et al.*, 1996; Love *et al.*, 1996; Shimizu *et al.*, 1996; Steinkuhler *et al.*, 1996*b*; Sudo *et al.*, 1996; Landro *et al.*, 1997), yeast (Song *et al.*, 1996; Markland *et al.*, 1997), recombinant baculovirus-infected insect cells (Overton *et al.*, 1995; Suzuki *et al.*, 1995; Steinkuhler *et al.*, 1996*a*; Zhang *et al.*, 1997), transient expression in mammalian cells (Bartenschlager *et al.*, 1993, 1994; Eckart *et al.*, 1993; Bouffard *et al.*, 1995; Failla *et al.*, 1995; Lin *et al.*, 1995; Reed *et al.*, 1995; Morgenstern *et al.*, 1997; Muramatsu *et al.*, 1997) and translation systems *in vitro* (Han *et al.*, 1995; Hahm *et al.*, 1995; Lin & Rice, 1995; Butkiewicz *et al.*, 1996; Pieroni *et al.*, 1997). NS3^{pro} expressed in such systems has been analysed using both polyprotein and peptidic substrates for the characterization of substrate specificities and to screen proteinase inhibitors.

Using NS3^{pro} expressed and purified from yeast together with hepatitis C virus protein substrates derived from transcription/translation of subgenomic cDNA clones *in vitro* Markland *et al.* (1997) identified a previously unreported cleavage site in the NS5A region (Thr^{2172} ⇓ Ser^{2173}), the cleavage occurring in the presence or absence of the NS4A activating peptide. Studies on the cleavage of synthetic peptide substrates have yielded much data on substrate specificities, kinetic parameters and inhibition (discussed in Clarke, 1997). The sensitivity of such *in vitro* analyses has been increased > 100-fold by the substitution of the scissile peptide bond by an ester linkage (Bianchi *et al.*, 1996). Analyses of the efficiency of cleavage (k_{cat}/K_m) of synthetic peptides corresponding to polyprotein sites cleaved in *trans* showed that sites were processed with quite different kinetics. Peptides corresponding to the NS4A/NS4B, 4B/5A and 5A/5B sites were cleaved by NS3^{pro} (activated by NS4A) at 92, 1130 and 8300 M^{-1} s^{-1}, respectively (Zhang *et al.*, 1997), and at 1600, 110 and 20000 M^{-1} s^{-1}, respectively, by Landro *et al.* (1997) – although the peptidic substrates used in the two studies were by no means identical.

The expression and purification of NS3-NS4A from COS-

7 cells permitted a comparison of the activity between NS3-4A and the proteolytic domain ($NS3^{pro}$) expressed separately (Morgenstern *et al*., 1997). Using NS4B-NS5A as a substrate the NS4B/NS5A cleavage was observed together with the cleavage between Thr^{2172} and Ser^{2173} reported by Markland *et al*. (1997). Intriguingly, the NS3-4A complex showed stimulation of its proteolytic, nucleoside triphosphatase and helicase activities by the addition of polynucleotides. The proteolytic activity was increased up to ~ 5-fold by the addition of poly(U), which could be abolished by the addition of salt (NaCl; 300 mM). This effect was not observed with the $NS3^{pro}$ domain either alone, nor with an NS4A activating peptide. The authors pointed out that although the mechanism of this stimulation is not clear, and may have a trivial explanation, the maximal stimulation of proteolytic activity [achieved with poly(U)] was paralleled by the polynucleotide binding specificity of the helicase domain.

An alternative approach to heterologous expression/purification of $NS3^{pro}$ has been the construction of 'surrogate' or chimeric viruses in which hepatitis virus NS3-4A sequences were inserted into picornavirus or alphavirus genomes in such a manner that viability of the rescued viruses depends upon $NS3^{pro}$ activity (Hahm *et al*., 1996; Filocamo *et al*., 1997). In both cases $NS3^{pro}$ sequences, together with an appropriate cleavage site, preceded the long open reading frame of the 'host' genomes. The effect of drugs may, therefore, be evaluated by the growth of these surrogate viruses. A 'genetic' screen for proteinase inhibitors has been developed in which the *lexA* DNA-binding domain is linked to the *gal*4 transcriptional domain via an $NS3^{pro}$ cleavage site (Song *et al*., 1996).

(iv) The NS2-3 proteinase

Analyses of the polyprotein processing activities of site-directed mutants of hepatitis C virus $NS3^{pro}$ showed that mutation of Ser^{139}, whilst abolishing proteolysis at known NS3 cleavage sites, did not affect cleavage of the NS2B/NS3 site. Deletion analyses showed this proteolytic activity to be a property of NS2 in combination with $NS3^{pro}$ (NS2-3; Grakoui *et al*., 1993*b*; Hijikata *et al*., 1993*a*). This NS2-3 proteinase activity is (weakly) inhibited by EDTA and stimulated by zinc (Grakoui *et al*., 1993*b*; Hijikata *et al*., 1993*b*; Reed *et al*., 1995; Kim *et al*., 1996; Love *et al*., 1996). Cleavage at the NS2/3 site of hepatitis C virus polyprotein was shown to be stimulated by the addition of microsomal membranes and resulted in the ultimate insertion of NS2 into the membranes (Pieroni *et al*., 1997). The NS2/3 cleavage was also stimulated by detergents, could be inhibited by alkylating agents or metal chelators (the latter inhibition could be reversed to some extent by the addition of $ZnCl_2$) and was sensitive to the redox state of the reaction mixture. A detailed discussion of this unusual cleavage can be found in Clarke (1997). Such experimental data are not available for hepatitis G virus NS3 but examination of a sequence alignment of the $NS3^{pro}$ regions of flavi-, pesti-, hepatitis C and G viruses (Fig. 2) shows that the histidine and three cysteines residues involved in zinc binding (see below) are conserved in both hepatitis C and G viruses, but not in flavi- or pestiviruses (Fig. 2).

(v) Pestivirus NS2/NS3 cleavage

Pestiviruses may exist in two forms or 'biotypes': cytopathogenic (CP) or non-cytopathogenic (non-CP). More specifically, viruses of one biotype derived from the other are regarded as 'virus pairs'. The mechanisms whereby viruses interconvert between biotypes is discussed in Becher *et al*. (1996) and Meyers & Thiel (1996) – further discussion on the consequences for polyprotein processing can be found in Xu *et al*. (1997). Briefly, insertion of cellular sequences (notably ubiquitin) into the NS2/NS3 (p125) region of the polyprotein, duplication of viral sequences into, or deletions from, this same region result in a CP biotype with altered polyprotein processing – in the case of ubiquitin insertions by creating a cleavage site for the cellular enzyme ubiquitin carboxyterminal hydrolase (Meyers *et al*., 1991). In this type of processing NS3 (p80) is generated; in the case of BVDV the presence of NS3 in infected cells is taken to be diagnostic of a CP biotype. This is not the case for other pestiviruses such as classical swine fever (CSFV) or border disease (BDV) viruses where some NS3 can be detected in non-CP biotype infected cells.

Comparison of a CP and non-CP BVD virus pair (strains NCP7 and CP7, respectively) revealed CP7 to bear a 27 nt insertion (maintaining the single, long ORF) in the N-terminal region of NS2 (Tautz *et al*., 1997). Polyprotein processing studies on the NCP7 and CP7 viruses showed the CP biotype to produce the NS3 cleavage product, whereas the NCP7 biotype did not. As in the case of hepatitis C virus, mutation of the BVDV $NS3^{pro}$ serine nucleophile to alanine did not abrogate BVDV NS2/NS3 cleavage (cytopathic strains CP7 and NADL: Tautz *et al*., 1997; Xu *et al*., 1997).

(vi) Atomic structure of the NS3 proteolytic domain

The numbering scheme used here refers to the hepatitis C virus 1a sequence shown in Fig. 2. A simple illustration is provided to show the relative locations of the features discussed below (Fig. 4). The atomic co-ordinates of an $NS3^{pro}$/NS4A complex have recently been made available on the databases – not for the structure of Kim *et al*. (1996) nor Love *et al*. (1996) but from another laboratory working with the BK virus strain (Yan *et al*., 1998). The $NS3^{pro}$/NS4A atomic co-ordinate (PDB) file 1JXP, together with KINEMAGE and RASMOL files constructed to show the salient features of the $NS3^{pro}$/4A structure, are available at the FTP site given in the Introduction.

The overall architecture of the hepatitis C virus $NS3^{pro}$ is that of two six-stranded β-barrels characteristic of the chymotrypsin-like fold (Kim *et al*., 1996; Love *et al*., 1996; Yan *et*

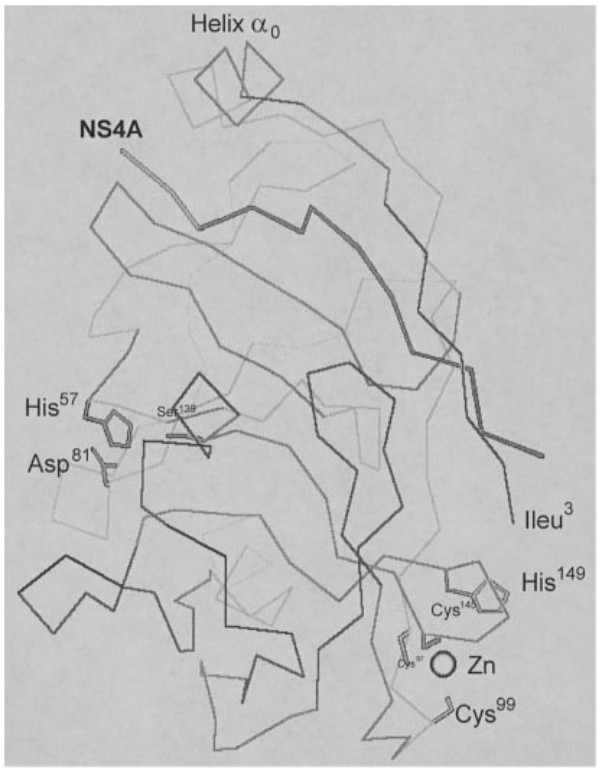

Fig. 4. Atomic structure of the NS3^{pro}/NS4A peptide. The α-carbon backbone of NS3^{pro} is shown (using atomic co-ordinates in PDB file 1JXP) together with the synthetic peptide corresponding to NS4A, the side-chains of the catalytic site residues, residues involved in binding zinc and the relative positions of the N terminus of NS3^{pro} and the zinc atom.

al., 1998). The availability of three structures, two solved in the presence of the activating NS4A peptide, one without, will permit some detailed comparisons – when all sets of atomic co-ordinates are available. The structure solved by Kim *et al*. was NS3^{pro} complexed with a synthetic oligopeptide, corresponding to residues Gly^{21}-Pro^{39} of the NS4A peptide, whilst the structure of Yan *et al*. was solved with a peptide corresponding to NS4A residues Gly^{21}-Arg^{34}. In each case the peptide shows an interaction with the extended N-terminal region and with the core of the enzyme. The interactions between NS3^{pro} and the (longer) NS4 peptide described in Kim *et al*. involve all but 2 out of 19 NS4A residue main-chain carbonyl or amide groups in hydrogen bonding. In addition interactions between NS4A hydrophobic side-chains and NS3 pockets constitute the hydrophobic core of the N-terminal region. The mutation of these buried residues proved to have the greatest affect on the ability of the NS4A peptide to activate the proteolytic activity. Deletions of NS3^{pro} N-terminal residues have shown that this region is important in NS3:NS4A interactions. The N-terminal 20 residues of NS4A are predicted to form a membrane-spanning helix and it has been postulated that this region serves either to anchor or target NS3, possibly as part of a larger replication complex, to membranes. In the structures of Kim *et al*. and Yan *et al*. the catalytic residues adopt the canonical catalytic triad steric pattern: the carboxyl group of Asp^{81} and the hydroxyl of Ser^{139} oriented towards the imidazole ring of His^{57}. In the case of the structure of Love *et al*., solved in the absence of the activating NS4A sequences, the side-chain of Asp^{81} is oriented away from His^{57} and forms an ion-pair with Arg^{155}. Alignments show that Arg^{155} is conserved amongst hepatitis C virus NS3 sequences and that there is a conserved lysine in pestivirus sequences, but no corresponding conserved basic residue in flavi- or hepatitis G virus sequences (Fig. 2). The orientation of the aspartate side-chain (away from the catalytic histidine imidazol ring) has been observed in the picornavirus hepatitis A 3C proteinase structure (Allaire *et al*., 1994; discussed in Ryan & Flint, 1997), although since the active site nucleophile was a cysteine in this case it was proposed that the proteolytic mechanism involved a catalytic dyad rather than a triad. Since the activation of NS3 by NS4A occurs in *trans*, active vs inactive enzyme is not a function of alternative protein folding pathways but a structural modification of NS3 by NS4A. It is not clear how such an interaction could result in activation of the enzyme by reorientation of the side-chain of Asp^{81} towards His^{57}.

The oxyanion hole (a series of pre-aligned dipoles interacting with the peptide bond carbonyl oxygen and thought to promote the formation of the tetrahedral transition state) is formed by the main-chain amides of Gly^{137} and Ser^{139}. Modelling of substrate in the binding pocket indicated favourable interactions of the P1 residue (Thr or Cys) hydroxyl or sulphydryl groups with the delocalized electron cloud of the aromatic ring of Phe^{154} located at the bottom of the P1 binding pocket (Pizzi *et al*., 1994; Kim *et al*., 1996; Love *et al*., 1996; Yan *et al*., 1998). Hepatitis C virus NS3^{pro} residues Leu^{135}, Ala^{157}, Arg^{161} and Lys^{165} (see Fig. 2) are thought to contribute to the binding of P2–P6 via interactions with substrate main-chain groups (with an apparent lack of interaction with side-chains) – a more extended *β*-interaction than is observed in most other proteinase structures.

NS3^{pro} binds zinc through a tetrahedral interaction with Cys^{97}, Cys^{99}, Cys^{145} and a water molecule within hydrogen-bonding distance of His^{149} (Kim *et al*., 1996; Love *et al*., 1996; Yan *et al*., 1998; see Figs 2 and 4). These residues are located in the turns of loop F1–A2, joining the domains, and the loop joining D2–E2. Since the zinc atom is some 20 Å from the active site it was assumed that the metal ion does not play an active role in (serine proteinase) catalysis, rather a structural one (see below).

(vii) Polyprotein cleavage sites

Understanding of polyprotein processing is more advanced for the flaviviruses than for other members of the *Flaviviridae*. A complex picture has emerged with a combination of host-cell- and virus-encoded proteinases processing at sites in both the structural and non-structural domains of the polyprotein (Fig. 1).

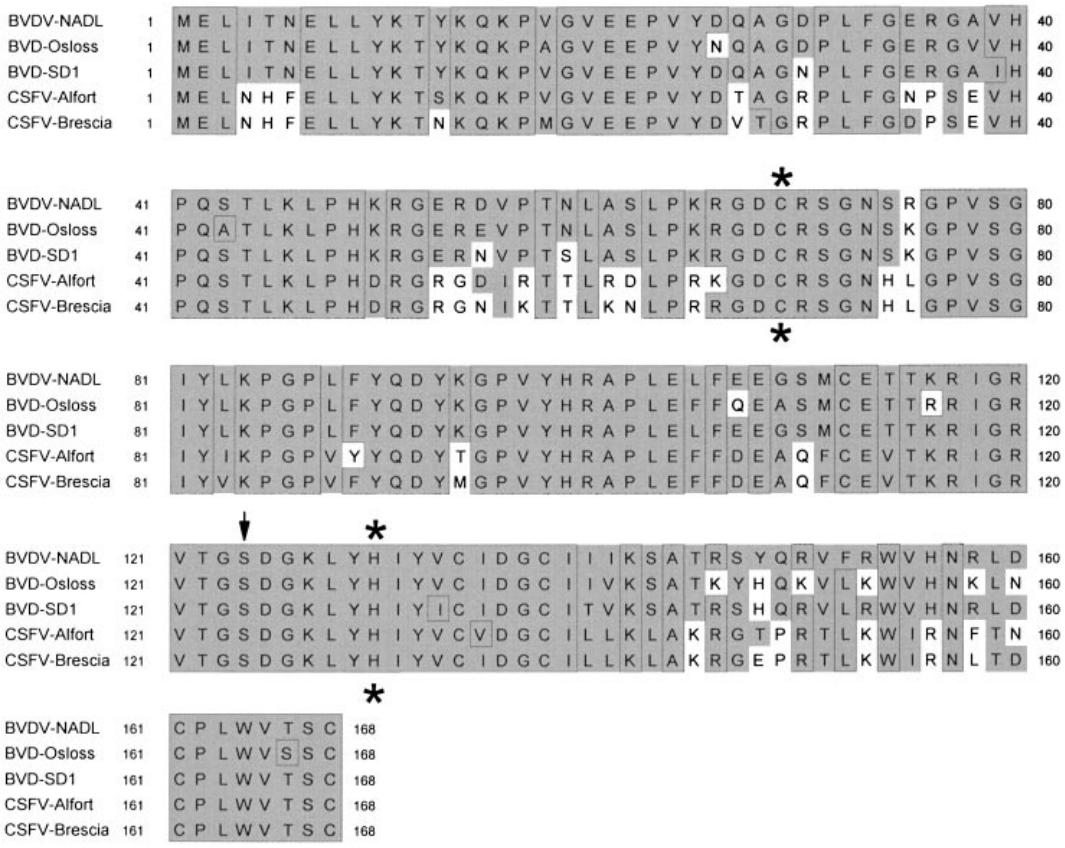

Fig. 5. Sequence alignment of pestivirus p20 proteinases. The N-terminal regions of pestivirus polyproteins were aligned using CLUSTALW. The position of the $Ser^{124} \rightarrow$ Ala mutated in Wiskerchen *et al.* (1991) is shown (arrow) and the putative active site nucleophiles predicted in Stark *et al.* (1993) (asterisks).

(a) **Flaviviruses.** Experiments on a number of different flaviviruses showed that the cleavages between the C/prM, prM/E, E/NS1, NS1/NS2A and NS4A-B proteins are mediated by host-cell proteinases (Biedrzycka *et al.*, 1987; Nowak *et al.*, 1989; Ruiz-Linares *et al.*, 1989; Speight *et al.*, 1988; Markoff, 1989; Wright *et al.*, 1989; Lin *et al.*, 1993; Lobigs, 1993; Yamshchikov & Compans, 1994, 1995; Falgout & Markoff, 1995). Sites of other processing events (identified by N-terminal sequencing of virus proteins) occurred at motifs commonly consisting of dibasic amino acid pairs at P1 and P2 followed by a small, non-branched amino acid at P1′ (Rice *et al.*, 1986; Biedrzycka *et al.*, 1987; Speight *et al.*, 1988; Chambers *et al.*, 1989). These sites proved to be conserved amongst the flavivirus sequences then becoming available (Rice *et al.*, 1985; Castle *et al.*, 1986; Sumiyoshi *et al.*, 1987; Coia *et al.*, 1988; Hahn *et al.*, 1988; Pletnev *et al.*, 1990). Recently, it has been shown that a cleavage within the final helicase motif in the C-terminal region of dengue virus type 2 NS3 (Fig. 1) is mediated by the NS2B/3 proteinase (mol. mass 69 kDa; see below) producing the cleavage products NS3′ (50 kDa) and NS3″ (19 kDa; Teo & Wright, 1997), although the role of this cleavage in virus replication is not known.

(b) **Pestiviruses.** Recently, the non-structural protein cleavage sites 3/4A, 4A/4B, 4B/5A and 5A/5B have been determined by N-terminal sequencing of processing products and, in addition, shown to be cleaved by the $NS3^{pro}$ serine-type proteinase activity (Tautz *et al.*, 1997; Xu *et al.*, 1997). The consensus cleavage site specificities were determined as leucine at P1 and either serine, alanine or asparagine at P1′.

(c) **Hepatitis C virus.** The cleavage site specificities have been discussed elsewhere (Mills, 1996; Clarke, 1997). In summary, the substrate binding specificities of $NS3^{pro}$ are: (i) an acidic residue at P6, (ii) cysteine at P1 (cleavage in *trans*), (iii) threonine at P1 (cleavage in *cis* at the NS3/NS4A site) and (iv) serine or alanine at P1′. The unusual NS2/NS3 cleavage event described above has been reviewed recently (Clarke, 1997).

(d) **Hepatitis G virus.** The putative NS3 proteinase domain of hepatitis GB virus B has been expressed and its activity tested with hepatitis C virus substrates (Scarselli *et al.*, 1997). The proteinase was able to cleave at the hepatitis C virus NS4A/NS4B, NS4B/NS5A and NS5A/NS5B sites but was not activated by a peptide (corresponding to residues 21–34 of

hepatitis C virus NS4A) shown to activate hepatitis C virus $NS3^{pro}$. By alignment with hepatitis C virus cleavage sites the hepatitis GB virus B $NS3^{pro}$ appears to have a specificity (*trans* cleavage) similar to hepatitis C virus $NS3^{pro}$ – cysteine at P1 and alanine, glycine or serine at P1′.

Pestivirus N-terminal proteinase

Expression of subgenomic cDNA clones derived from the N-terminal region of the BVDV polyprotein showed a 20 kDa product produced by autocatalytic cleavage of the N-terminal proteinase (N^{pro}; Wiskerchen *et al.*, 1991; Muyldermans *et al.*, 1997). Analysis of sequence alignments suggested that other viruses had proteins located at the N terminus of their polyproteins with similarities to papain-related cysteine-type proteinases (Gorbalenya *et al.*, 1991). This has been shown to be the case for a number of viruses within the picornavirus super-group (reviewed by Ryan & Flint, 1997). Interestingly, one type of insertion into the NS2/NS3 (p125) region of the polyprotein involves a duplication of a proteolytically active N^{pro} (Meyers *et al.*, 1992). The N terminus of NS3 is generated in this case by the autocatalytic N^{pro} C-terminal cleavage activity – producing a cytopathogenic virus. Characterization of the pestivirus N^{pro} is incomplete although analysis of polyprotein processing of the N^{pro} duplication form indicates the C-terminal cleavage site as shown in Fig. 5 (Stark *et al.*, 1993). Identification of the active site residues by site-directed mutagenesis is incomplete but alignments suggest N^{pro} to be a thiol-type proteinase with Cys^{69} and His^{130} to be good candidates for active site residues (Stark *et al.*, 1993; Fig. 5).

The role of such N-terminal proteinases could be either to generate an N terminus (of the adjacent translation product) that is more suitable for a specific post-translational modification, to cleave (in *trans*) at other sites within the polyprotein or to cleave a cellular protein(s), as is the case for foot-and-mouth disease virus L^{pro} (Devaney *et al.*, 1988).

Concluding remarks

The N-terminal proteolytic domain of the NS3 protein of viruses within the *Flaviviridae* is a remarkably versatile enzyme – in the sense that it appears to be able to be regulated by a quite a variety of oligopeptide sequences. Positive regulation (activation) is observed by its interaction with NS2B (flaviviruses), or NS4A (pesti- and hepatitis C viruses). It is reasonable to assume that the same type of intimate molecular association observed between hepatitis C virus $NS3^{pro}$ and NS4A is also present in pesti- and hepatitis GB viruses. The residues involved in the binding of zinc by hepatitis C virus $NS3^{pro}$ are conserved in all hepatitis viruses, but not flavi- or pestiviruses (Fig. 2). Although there is no suggestion that the zinc atom in the atomic structure plays a role in the serine proteinase activity of $NS3^{pro}$, it is interesting to note that zinc, when found as a structural component of proteins, is bound by four cysteine residues whereas zinc in proteins playing a catalytic role is bound by three residues and an activated water molecule – as is thought to be case for hepatitis C virus $NS3^{pro}$. This may suggest that NS2-3 protein of the hepatitis viruses is a metallo-proteinase cleaving at the NS2/NS3 site.

The α-carbon of Ile^{3} (the first residue for which electron density is observed in the structure of Yan *et al.*, 1998) is some 13 Å from the zinc atom. We have modelled the presence of the two N-terminal residues (Ala-Pro) onto this structure. Although the loop joining sheets D2 and E2 lies between the N terminus of NS3 and the zinc atom (Fig. 4) the A_0 β-strand (lying on the surface of the molecule) can be re-modelled by altering bond angles either side of helix α_0 to bring the N terminus into an appropriate sterochemistry for nucleophilic attack by the zinc atom. It is noteworthy that the activating NS4A peptide is not required for the NS2/NS3 cleavage and that 'removing' this feature from the structure allows the A_0 strand greater freedom in this respect.

The NS3 cleavage product of this hypothetical metallo-proteinase activity would, itself, possess *serine* proteinase activity. In a sense this would be the molecular antithesis of an enzyme such as wheat serine carboxypeptidase II in which the active site nucleophiles are present on two chains (chain a – Ser^{146}; chain b – Asp^{338} and His^{397}) in a metallo-proteinase-type fold (Liao & Remington, 1990). In the case of the hepatitis C virus NS2-3 protein the putative metallo-proteinase would (at least in part) be composed of a serine proteinase-type fold.

The model outlined above is, in many ways, an attractive one and consistent with the lack of conservation of the zinc-binding residues in the pestiviruses. The NS2/3 protein of BVDV (the argument would run) does not possess metallo-proteinase activity and requires the insertion of cellular sequences such that a cleavage in this region of the polyprotein is brought about by a cellular enzyme, or by the insertion of a *cis*-acting virus proteinase (N^{pro}). The 'acquisition' of a CP biotype and NS2/NS3 polyprotein cleavage by the insertion of just 9 aa in the NS2 region (Tautz *et al.*, 1997) obviously requires further work to determine if this event has conferred some form of cellular proteinase cleavage site. All appears well until one considers how the other pestiviruses, CSFV or BDV, process their polyproteins to produce NS3? The atomic structures of hepatitis C virus NS2-3 protein and the 40 aa tract of the $NS2B/NS3^{pro}$ complex of flaviviruses are eagerly awaited!

References

Adams, A., Chamberlain, R. W., Taylor, L. A., Davidson, F., Lin, C. K., Simmonds, P. & Elliott, R. M. (1997). Complete coding sequence of hepatitis C virus genotype 6a. *Biochemical and Biophysical Research Communications* **234**, 393–396.

Allaire, M., Cheraia, M. M., Malcolm, B. A. & James, M. N. G. (1994). Picornaviral 3C cysteine proteinases have a fold similar to chymotrypsin-like serine proteinases. *Nature* **369**, 72–76.

Arias, C., 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. L., Ahlborn-Laake, L., Mous, J. & Jacobsen, H. (1993). Nonstructural protein 3 of the hepatitis C virus encodes a serine-type proteinase required for cleavage at the NS3/4 and NS4/5 junctions. *Journal of Virology* **67**, 3835–3844.

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

Bartenschlager, R. L., Lohmann, V., Wilkinson, T. & Koch, J. O. (1995). Complex formation between the NS3 serine-type proteinase of the hepatitis C virus and NS4A and its importance for polyprotein maturation. *Journal of Virology* **69**, 7519–7528.

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.

Becher, P., Meyers, G., Shannon, A. D. & Thiel, H.-J. (1996). Cytopathogenicity of border disease virus is correlated with integration of cellular sequences into the viral genome. *Journal of Virology* **70**, 2992–2998.

Bianchi, E., Steinkuhler, C., Taliani, M., Urbani, A., De Francesco, R. & Pessi, A. (1996). Synthetic depsipeptide substrates for the assay of human hepatitis C virus protease. *Analytical Biochemistry* **237**, 239–244.

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.

Bouffard, P., Bartenschlager, R., Ahlborn-Laake, L., Mous, J., Roberts, N. & Jacobsen, H. (1995). An *in vitro* assay for hepatitis C virus serine proteinase. *Virology* **209**, 52–59.

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., Ramanthan, 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.

Cahour, A., Falgout, B. & Lai, C.-J. (1992). Cleavage of the dengue virus polyprotein at the NS3/NS4A and NS4B/NS5 junctions is mediated by viral protease NS2B-NS3, whereas NS4A/NS4B may be processed by a cellular protease. *Journal of Virology* **66**, 1535–1542.

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

Chamberlain, R. W., Adams, N., Saeed, A. A., Simmonds, P. & Elliott, R. M. (1997*a*). Complete nucleotide sequence of a type 4 hepatitis C virus variant, the predominant genotype in the Middle East. *Journal of General Virology* **78**, 1341–1347.

Chamberlain, R. W., Adams, N. J., Taylor, L. A., Simmonds, P. & Elliott, R. M. (1997*b*). The complete coding sequence of hepatitis C virus genotype 5a, the predominant genotype in South Africa. *Biochemical and Biophysical Research Communications* **236**, 44–49.

Chambers, T. J., McCourt, D. W. & Rice, C. M. (1989). Yellow fever virus proteins NS2a, NS2b and NS4b; identification and partial amino acid sequence analysis. *Virology* **169**, 100–109.

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

Chambers, T. J., Wier, 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 yellow fever virus NS3 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., Grakoui, A. & Rice, C. M. (1991). Processing of the yellow fever virus nonstructural polyprotein: a catalytically active NS3 protease domain and NS2B are required for cleavages at dibasic sites. *Journal of Virology* **65**, 6042–6050.

Chayama, K., Tsubota, A., Koida, I., Arase, Y., Saitoh, S., Ikeda, K. & Kumada, H. (1994). Nucleotide sequence of hepatitis C virus (type 3b) isolated from a Japanese patient with chronic hepatitis C. *Journal of General Virology* **75**, 3623–3628.

Choo, Q. L., Richman, K., Han, J. H., Berger, K., Lee, C., Dong, C., Gallegos, C., Coit, D., Medina-Selby, A., Barr, P. J., Weiner, A., Bradley, D. W., Kuo, G. & Houghton, M. (1991). Genomic organization and diversity of the hepatitis C virus. *Proceedings of the National Academy of Sciences, USA* **88**, 2451–2455.

Clarke, B. (1997). Molecular virology of hepatitis C virus. *Journal of General Virology* **78**, 2397–2410.

Cleaves, G. R. (1985). Identification of dengue type 2 virus-specific high molecular weight proteins in virus-infected BHK cells. *Journal of General Virology* **66**, 2767–2771.

Coia, G., Parker, M. D., Speight, G., Byrne, M. E. & Westaway, E. G. (1988). Nucleotide and complete amino acid sequences of Kunjin virus: definitive gene order and characteristics of the virus-specified proteins. *Journal of General Virology* **69**, 1–21.

Collett, M. S., Larson, R., Belzer, S. K. & Retzel, E. (1988*a*). Pestivirus p80 (p125) may be a bifunctional protein with helicase and protease activity. *Virology* **165**, 200–208.

Collet, M. S., Larson, R., Gold, C., Strick, D., Anderson, D. K. & Purchio, A. F. (1988*b*). Molecular cloning and nucleotide sequence of the pestivirus bovine viral diarrhea virus. *Virology* **165**, 191–199.

Dalgarno, L., Trent, D. W., Strauss, J. H. & Rice, C. M. (1986). Partial nucleotide sequence of the Murray Valley encephalitis virus genome – comparison of the encoded polypeptides with yellow fever virus structural and non-structural proteins. *Journal of Molecular Biology* **187**, 309–323.

De Moerlooze, L., Lecomte, C., Brown-Shimmer, S., Schmetz, D., Guiot, C., Vandenbergh, D., Allaer, D., Rossius, M., Chappuis, G., Dina, D., Renard, A. & Martial, J. A. (1993). Nucleotide sequence of the bovine viral diarrhoea virus Osloss strain: comparison with related viruses and identification of specific DNA probes in the 5′ untranslated region. *Journal of General Virology* **74**, 1433–1438.

Deng, R. & Brock, K. V. (1992). Molecular cloning and nucleotide sequence of a pestivirus genome, noncytopathic bovine viral diarrhea virus strain SD-1. *Virology* **191**, 867–869.

Devaney, M. A., Vakharia, V. N., Lloyd, R. E., Ehrenfeld, E. & Grubman, M. J. (1988). Leader protein of foot-and-mouth disease virus is required for cleavage of the p220 component of the cap-binding protein complex. *Journal of Virology* **62**, 4407–4409.

Ding, J. Z., McGrath, W. J., Sweet, R. M. & Mangel, W. F. (1996). Crystal-structure of the human adenovirus proteinase with its 11-amino-acid cofactor. *EMBO Journal* **15**, 1778–1783.

Dougherty, W. G. & Semler, B. L. (1993). Expression of virus-encoded proteinases: functional and structural similarities with cellular enzymes. *Microbiological Reviews* **57**, 781–822.

D'Souza, E. D. A., Grace, K., Sangar, D. V., Rowlands, D. J. & Clarke, B. E. (1995). *In vitro* cleavage of hepatitis C virus polyprotein substrates by purified recombinant NS3 protease. *Journal of Virology* **76**, 1729–1736.

Eckart, M. R., Selby, M., Masiarz, F., Lee, C., Berger, K., Crawford, K., Kuo, C., Kuo, G., Houghton, M. & Choo, Q.-L. (1993). The hepatitis C

virus encodes a serine proteinase involved in processing of the putative nonstructural proteins from the viral polyprotein precursor. *Biochemical and Biophysical Research Communications* **192**, 399–406.

Failla, C., Tomei, L. & De Francesco, R. (1994). Both NS3 and NS4A are required for proteolytic processing of hepatitis C virus nonstructural proteins. *Journal of Virology* **68**, 3753–3760.

Failla, C., Tomei, L. & De Francesco, R. (1995). An amino-terminal domain of the hepatitis C virus NS3 protease is essential for interaction with NS4A. *Journal of Virology* **69**, 1769–1777.

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., Channock, R. & Lai, C.-J. (1989). Proper processing of dengue virus nonstructural glycoprotein NS1 requires the N-terminal hydrophobic signal sequence and the downstream nonstructural protein NS2a. *Journal of Virology* **63**, 1852–1860.

Falgout, B., Pethel, 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.

Felsenstein, J. (1991). PHYLIP manual, version 3.4, University Herbarium, University of California, Berkeley, California.

Filocamo, G., Pacini, L. & Migliaccio, G. (1997). Chimeric Sindbis viruses dependent on the NS3 protease of hepatitis C virus. *Journal of Virology* **71**, 1417–1427.

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., Blinov, V. M. & Koonin, E. V. (1989*a*). Cysteine proteases of positive strand RNA viruses and chymotrypsin-like serine proteases. A distinct protein superfamily with a common structural fold. *FEBS Letters* **243**, 103–114.

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

Gorbalenya, A. G., Koonin, E. V. & Lai, M. M.-C. (1991). Putative papain-related thiol proteases of positive strand RNA viruses. *FEBS Letters* **288**, 201–205.

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

Grakoui, A., McCourt, D. W., Wychowski, C., Feinstone, S. M. & Rice, C. M. (1993*b*). A second hepatitis C virus-encoded serine proteinase. *Proceedings of the National Academy of Sciences, USA* **90**, 10583–10587.

Hahm, B., Han, D. S., Back, S. H., Song, O.-K., Cho, M.-J., Kim, C.-J., Shimotohno, K. & Jang, S. K. (1995). NS3-4A of hepatitis C virus is a chymotrypsin-like protease. *Journal of Virology* **69**, 2534–2539.

Hahm, B., Back, S. H., Lee, T. G., Wimmer, E. & Jang, S. K. (1996). Generation of a novel poliovirus with a requirement of hepatitis C virus NS3 activity. *Virology* **226**, 318–326.

Hahn, Y. S., Galler, R., Hunkapiller, T., Dalrymple, J. M., Strauss, J. H. & Strauss, E. G. (1988). Nucleotide sequence of dengue 2 RNA and comparison of the encoded proteins with those of other flaviviruses. *Virology* **162**, 167–180.

Han, D. S., Hahm, B., Rho, H.-M. & Jang, S. K. (1995). Identification of the protease domain in NS3 of hepatitis C virus. *Journal of General Virology* **76**, 985–993.

Higgins, D. E., Bleasby, A. J. & Fuchs, R. (1991). Improved software for multiple sequence alignment. *Computer Applications in the BioSciences* **8**, 189–191.

Hijikata, M., Mizushima, H., Tanji, Y., Komoda, Y., Hirowatari, Y., Akagi, T., Kato, N., Kimura, K. & Shimotohno, K. (1993*a*). Two distinct proteinase activities required for the processing of a putative non-structural precursor protein of hepatitis C virus. *Journal of Virology* **67**, 4665–4675.

Hijikata, M., Mizushima, H., Akagi, T., Mori, S., Kakiuchi, N., Kato, N., Tanaka, T., Kimura, K. & Shimotohno, K. (1993*b*). Proteolytic processing and membrane association of putative nonstructural proteins of hepatitis C virus. *Proceedings of the National Academy of Sciences, USA* **90**, 10773–10777.

Kakiuchi, N., Hijikata, M., Komoda, Y., Tanji, Y., Hirowatari, Y. & Shimotohno, K. (1995). Bacterial expression and analysis of cleavage activity of HCV serine proteinase using recombinant and synthetic substrate. *Biochemical and Biophysical Research Communications* **210**, 1059–1065.

Kato, N., Hijikata, M., Ootsuyama, Y., Nakagawa, M., Ohkoshi, S., Sugimura, T. & Shimotohno, K. (1990). Molecular cloning of the human hepatitis C virus genome from Japanese patients with non-A, non-B hepatitis. *Proceedings of the National Academy of Sciences, USA* **87**, 9524–9528.

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.

Kim, D. W., Gwack, Y., Han, J. H. & Choe, J. (1997). Towards defining a minimal functional domain for NTPase and RNA helicase activities of the hepatitis C virus NS3 protein. *Virus Research* **49**, 17–25.

Koch, J. O., Lohmann, V., Herian, U. & Bartenschlager, R. (1996). In vitro studies on the activation of the hepatitis C virus NS3 proteinase by the NS4 cofactor. *Virology* **221**, 54–66.

Komoda, Y., Hijikata, M., Sato, S., Shin-Ichi, A., Kimura, K. & Shimotohno, K. (1994). Substrate requirements of hepatitis C virus serine proteinase for intermolecular polypeptide cleavage in *Escherichia coli*. *Journal of Virology* **68**, 7351–7357.

Koonin, E. V. & Dolja, V. V. (1993). Evolution and taxonomy of positive strand RNA viruses: implications of comparative analysis of amino acid sequences. *Critical Reviews in Biochemistry and Molecular Biology* **28**, 375–430.

Landro, J. A., Raybuck, S. A., Luong, Y. P. C., O'Malley, E. T., Harbeson, S. L., Morgenstern, K. A., Rao, G. & Livingston, D. J. (1997). Mechanistic role of an NS4A peptide cofactor with the truncated NS3 protease of hepatitis C virus: elucidation of the NS4A stimulatory effect *via* kinetic analysis and inhibitor mapping. *Biochemistry* **36**, 9340–9348.

Leary, T. P., Muerhoff, A. S., Simons, J. N., Pilot-Matias, T. J., Erker, J. C., Chalmers, M. L., Schlauder, G. G., Dawson, G. J., Desai, S. M. & Mushahwar, I. K. (1995). Sequence and genomic organization of GBV-C: a novel member of the *Flaviviridae* associated with human non A–E hepatitis. *Journal of Medical Virology* **48**, 60–67.

Liao, D.-I. & Remington, S. J. (1990). Structure of wheat serine carboxypeptidase II at 3. 5 Å resolution. *Journal of Biological Chemistry* **265**, 6528–6531.

Lin, C. & Rice, C. M. (1995). The hepatitis C virus NS3 serine proteinase

and NS4A protein cofactor: establishment of a cell-free trans-processing assay. *Proceedings of the National Academy of Sciences, USA* **92**, 7622–7626.

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/B 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.

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

Manabe, S., Fuke, I., Tanishita, O., Kaji, C., Gomi, Y., Yoshida, S., Mori, C., Takamizawa, A., Yosida, I. & Okayama, H. (1994). Production of nonstructural proteins of hepatitis C virus requires a putative viral protease encoded by NS3. *Virology* **198**, 636–644.

Mangel, W. F., Toledo, D. L., Brown, M. T., Martin, J. H. & McGrath, W. J. (1996). Characterization of 3 components of human adenovirus proteinase activity in vitro. *Journal of Biological Chemistry* **271**, 536–543.

Markland, W., Petrillo, R. A., Fitzgibbon, M., Fox, T., McCarrick, R., McQuaid, T., Fulghum, J. R., Chen, W., Fleming, M. A., Thomson, J. A. & Chambers, S. P. (1997). Purification and characterization of the NS3 serine protease domain of hepatitis C virus expressed in *Saccharomyces cerevisiae*. *Journal of General Virology* **78**, 39–43.

Markoff, L. (1989). In vitro processing of dengue virus structural proteins: cleavage of the pre-membrane protein. *Journal of Virology* **63**, 3345–3352.

Meyers, G. & Thiel, H.-J. (1996). Molecular characterization of pestiviruses. *Advances in Virus Research* **47**, 53–118.

Meyers, G., Rumenapf, T. & Thiel, H.-J. (1989). Molecular cloning and nucleotide sequence of the hog cholera virus. *Virology* **171**, 555–567.

Meyers, G., Tautz, N., Dubovi, E. J. & Thiel, H.-J. (1991). Viral cytopathogenicity correlated with integration of ubiquitin-coding sequences. *Virology* **180**, 602–616.

Meyers, G., Tautz, N., Stark, R., Brownlie, J., Dubovi, E. J., Collet, M. S. & Thiel, H.-J. (1992). Rearrangement of viral sequences in cytopathogenic pestiviruses. *Virology* **191**, 368–386.

Meyers, G., Tautz, N., Becher, P., Thiel, H.-J. & Kuemmerer, B. M. (1996). Recovery of cytopathogenic and noncytopathogenic bovine viral diarrhea viruses from cDNA constructs. *Journal of Virology* **70**, 8606–8613.

Miller, R. H. & Purcell, R. H. (1990). Hepatitis C virus shares amino acid sequence similarity with pesti- and flaviviruses as well as members of two plant virus supergroups. *Proceedings of the National Academy of Sciences, USA* **87**, 2057–2061.

Mills, J. S. (1996). Virus proteinase inhibitors – what next after HIV? *Antiviral Chemistry & Chemotherapy* **7**, 281–293.

Moorman, R. J. M., Warmerdam, P. A. M., Van Der Meer, B., Schapper, W. M. M., Wensvoort, G. & Hulst, M. M. (1990). Molecular cloning and nucleotide sequence of hog cholera virus strain Brescia and mapping of the genomic region encoding envelope protein E1. *Virology* **177**, 184–198.

Moormann, R. J. M., van Gennip, H. G. P., Miedema, G. K. W., Hulst, M. M. & van Rijn, P. A. (1996). Infectious RNA transcribed from an engineered full-length cDNA template of the genome of a pestivirus. *Journal of Virology* **70**, 763–770.

Morgenstern, K. A., Landro, J. A., Hsiao, K., Lin, C., Gu, Y., Su, M. S.-S. & Thomson, J. A. (1997). Polynucleotide modulation of the protease, nucleoside triphosphatase and helicase activities of a hepatitis NS3-NS4A complex isolated from transfected COS cells. *Journal of Virology* **71**, 3767–3775.

Muramatsu, S., Ishido, S., Fujita, T., Itoh, M. & Hotta, H. (1997). Nuclear localization of the NS3 protein of hepatitis C virus and factors affecting the localization. *Journal of Virology* **71**, 4954–4961.

Muyldermans, G., SanGabriel, M. A. C., Hamers, R. & Wyns, L. (1997). Expression in *E. coli* and purification of the active autoprotease p20 of classical swine fever virus. *Virus Genes* **13**, 135–142.

Nakao, H., Okamoto, H., Tokita, H., Inoue, T., Iizuka, H., Pozzato, G. & Mishiro, S. (1996). Full-length genomic sequence of a hepatitis C virus genotype 2c isolate (BEBE1) and the 2c-specific PCR primers. *Archives of Virology* **141**, 701–704.

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.

Ohba, K.-I., Mizokami, M., Lau, J. Y. N., Orito, E., Ikeo, K. & Gojobori, T. (1996). Evolutionary relationship of hepatitis C, pesti-, flavi-, plant viruses and newly discovered GB hepatitis agents. *FEBS Letters* **378**, 232–234.

Okamoto, H., Okada, S., Sugiyama, Y., Kurai, K., Iizuka, H., Machida, A., Miyakawa, Y. & Mayumi, M. (1991). Nucleotide sequences of the genomic RNA of hepatitis C virus isolated from a human carrier: comparison with reported isolates for conserved and divergent regions. *Journal of General Virology* **72**, 2697–2704.

Okamoto, H., Kurai, K., Okada, S.-I., Yamamoto, K., Lizuka, H., Tanaka, T., Fukuda, S., Tsuda, F. & Mishiro, S. (1992). Full-length sequence of a hepatitis C virus genome having poor homology to reported isolates: comparative study of four distinct genotypes. *Virology* **188**, 331–341.

Okamoto, H., Kojima, M., Sakamoto, M., Iizuka, H., Hadiwandowo, S., Suwignyo, S., Miyakawa, Y. & Mayumi, M. (1994). The entire nucleotide sequence and classification of a hepatitis C virus isolate of a novel genotype from an Indonesian patient with chronic liver disease. *Journal of General Virology* **75**, 629–635.

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

Overton, H., McMillan, D., Gillespie, F. & Mills, J. (1995). Recombinant baculovirus-expressed NS3 proteinase of hepatitis C virus shows activity in cell-based and *in vitro* assays. *Journal of General Virology* **76**, 3009–3019.

Pieroni, L., Santolini, E., Fipaldini, C., Pacini, L., Migliaccio, G. & La Monica, N. (1997). In vitro study of the NS2-3 protease of hepatitis C virus. *Journal of Virology* **71**, 6373–6380.

Pizzi, E., Tramontano, A., Tomei, L., La Monica, N., Failla, C., Sardana, M., Wood, T. & De Francesco, R. (1994). Molecular model of the specificity pocket of the hepatitis C virus protease: implications for substrate recognition. *Proceedings of the National Academy of Sciences, USA* **91**, 888–892.

Pletnev, A. G., Yamshchikov, V. F. & Blinov, V. M. (1990). Nucleotide sequence of the genome and complete amino acid sequence of the polyprotein of tick-borne encephalitis virus. *Virology* **174**, 250–263.

Preugschat, F. & Strauss, J. H. (1991). Processing of nonstructural protein-NS4A and protein-NS4B of dengue-2 virus in vitro and in vivo. *Virology* **185**, 689–697.

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.

Pugachev, K. V., Nomokonova, N. Y., Dobrikova, E. Y. & Wolf, Y. I. (1993). Site-directed mutagenesis of the tick-borne encephalitis virus NS3 gene reveals the putative serine protease domain of the NS3 protein. *FEBS Letters* **328**, 115–118.

Reed, K. E., Grakoui, A. & Rice, C. M. (1995). Hepatitis C virus-encoded NS2-3 protease: cleavage site mutagenesis and requirements for bimolecular cleavage. *Journal of Virology* **69**, 4127–4136.

Renard, A., Dino, D. & Martial, J. (1987). Vaccines and diagnostics derived from bovine diarrhea virus. European Patent Application 86870095. 6. Publication no. 0208672 14 January 1987.

Rice, C. M. & Strauss, J. H. (1990). Production of flavivirus polypeptides by proteolytic processing. *Seminars in Virology* **1**, 357–367.

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.

Ridpath, J. F. & Bolin, S. R. (1995). The genomic sequence of a virulent bovine viral diarrhea virus (BVDV) from the type 2 genotype: detection of a large genomic insertion in a noncytopathic BVDV. *Virology* **212**, 39–46.

Ridpath, J. F., Bolin, S. R., Evermann, J. F., Landgraf, J. G. & Frey, M. L. The complete genomic sequence of the pestivirus border disease virus: comparison to genomic sequences of other pestiviruses. Unpublished. Accession no. U70263.

Roath, P. D. & Berry, E. S. Nucleotide sequence of ILLC-bovine viral diarrhea virus. Unpublished. Acc. No. U86599.

Ruiz-Linares, A., Cahour, A., Despres, P., Girard, M. & Bouloy, M. (1989). Processing of yellow fever virus polyprotein: role of cellular proteases in maturation of the structural proteins. *Journal of Virology* **63**, 4199–4209.

Ryan, M. D. & Flint, M. (1997). Virus-encoded proteinases of the picornavirus super-group. *Journal of General Virology* **78**, 699–723.

Sakamoto, M., Akahane, Y., Tsuda, F., Tanaka, T., Woodfield, D. G. & Okamoto, H. (1994). Entire nucleotide sequence and characterization of a hepatitis C virus of genotype V/3a. *Journal of General Virology* **75**, 1761–1768.

Satoh, S., Tanji, Y., Hijikata, M., Kimura, K. & Shimotohono, K. (1995). The N-terminal region of hepatitis C virus nonstructural protein 3 (NS3) is essential for stable complex formation with NS4A. *Journal of Virology* **69**, 4255–4260.

Scarselli, E., Urbani, A., Sbardellati, A., Tomei, L., De Francesco, R. & Traboni, C. (1997). GB virus B and hepatitis C virus NS3 serine proteases share substrate specificity. *Journal of Virology* **71**, 4985–4989.

Shimizu, Y., Yamaji, K., Masuho, Y., Yokota, T., Inoue, H., Sudo, K., Satoh, S. & Shimotohno, K. (1996). Identification of the sequence on NS4A required for enhanced cleavage of the NS4A/5B site by hepatitis C virus NS3 proteinase. *Journal of Virology* **70**, 127–132.

Simons, J. N., Pilot-Matias, T. J., Leary, T. P., Dawson, G. J., Desai, S. M., Schlauder, G. G., Muerhoff, A. S., Erker, J. C., Buijk, S. L., Chalmers, M. L., van Sant, C. L. & Mushahwar, I. K. (1995). Identification of two flavivirus-like genomes in the GB hepatitis agent. *Proceedings of the National Academy of Sciences, USA* **92**, 3401–3405.

Song, O.-K., Cho, O. H., Hahm, B. & Jang, S. K. (1996). Development of an *in vitro* assay system suitable for screening inhibitors of hepatitis C viral protease. *Molecules and Cells* **6**, 183–189.

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.

Stark, R., Meyers, G., Rumenapf, T. & Thiel, H.-J. (1993). Processing of pestivirus polyprotein: cleavage site between autoprotease and nucleocapsid protein of classical swine fever virus. *Journal of Virology* **67**, 7088–7095.

Steinkuhler, C., Tomei, L. & De Francesco, R. (1996*a*). *In vitro* activity of hepatitis C virus protease NS3 purified from recombinant baculovirus-infected Sf9 cells. *Journal of Biological Chemistry* **271**, 6367–6373.

Steinkuhler, C., Urbani, A., Tomei, L., Biasiol, G., Sardana, M., Bianchi, E., Pessi, A. & De Francesco, R. (1996*b*). Activity of purified hepatitis C virus protease NS3 on peptide substrates. *Journal of Virology* **70**, 6694–6700.

Sudo, K., Inoue, H., Shimizu, Y., Yamaji, K., Konno, K., Shigeta, S., Kaneko, T., Yokota, T. & Shimotohno, K. (1996). Establishment of an in vitro assay system for screening hepatitis C virus protease inhibitors using high performance liquid chromatography. *Antiviral Research* **32**, 9–18.

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.

Suzich, J. A., Tamura, J. K., Palmer-Hill, F., Warrener, P., Grakoui, A., Rice, C. M., Feinstone, S. M. & Collett, M. S. (1993). Hepatitis C virus NS3 protein polynucleotide-stimulated nucleoside triphosphatase and comparison with the related pestivirus and flavivirus enzymes. *Journal of Virology* **67**, 6152–6158.

Suzuki, T., Sato, M., Cheida, S., Shji, I., Harada, T., Yamakawa, Y., Watabe, S., Matsuura, Y. & Miyamura, T. (1995). *In vitro* and *in vivo trans*-cleavage activity of hepatitis C virus serine proteinase expressed by recombinant baculoviruses. *Journal of General Virology* **76**, 3021–3029.

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

Tautz, N., Elbers, K., Stoll, D., Meyers, G. & Thiel, H.-J. (1997). Serine protease of pestiviruses: determination of cleavage sites. *Journal of Virology* **71**, 5415–5422.

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.

Tomei, L., Failla, C., Santolini, E., De Francesco, R. & La Monica, N. (1993). NS3 is a serine protease required for processing of hepatitis C virus polyprotein. *Journal of Virology* **67**, 4017–4026.

Tomei, L., Failla, C., Vitale, R. L., Bianchi, E. & De Francesco, R. (1996). A central hydrophobic domain of the hepatitis C virus NS4A protein is necessary and sufficient for the activation of the NS3 protease. *Journal of General Virology* **77**, 1065–1070.

Warrerner, P. & Collet, M. S. (1995). Pestivirus NS3 (p80) protein possesses RNA helicase activity. *Journal of Virology* **69**, 1720–1726.

Webster, A., Hay, R. T. & Kemp, G. (1993). The adenovirus protease is activated by a virus-coded disulphide-linked peptide. *Cell* **72**, 97–104.

Wengler, G. & Wengler, G. (1991). The carboxy-terminal part of the NS3 protein of the West Nile flavivirus can be isolated as a soluble protein after proteolytic cleavage and represents an RNA-stimulated NTPase. *Virology* **184**, 707–715.

Wengler, G. & Wengler, G. (1993). The NS3 nonstructural protein of flavivirus contains an RNA triphosphatase activity. *Virology* **197**, 265–373.

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.

Wiskerchen, M. & Collett, M. S. (1991). Pestivirus gene-expression – protein p80 of bovine viral diarrhea virus is a proteinase involved in polyprotein processing. *Virology* **184**, 341–350.

Wiskerchen, M., Belzer, S. K. & Collet, M. S. (1991). Pestivirus gene expression: the first protein product of the bovine viral diarrhea virus large open reading frame, p20, possesses proteolytic activity. *Journal of Virology* **65**, 4508–4514.

Wright, P. S., Cauchi, M. R. & Ng, M. L. (1989). Definition of the carboxy termini of the three glycoproteins specified by dengue virus type 2. *Virology* **171**, 61–67.

Xu, J. A., Mendez, E., Caron, P. R., Lin, C., Murcko, M. A., Collet, M. S. & Rice, C. M. (1997). Bovine viral diarrhea virus NS3 serine proteinase: polyprotein cleavage sites, cofactor requirements, and molecular model of an enzyme essential for pestivirus replication. *Journal of Virology* **71**, 5312–5322.

Yamshchikov, V. F. & Compans, R. W. (1994). Processing of the intracellular form of the West Nile virus capsid protein by the viral NS2B-NS3 protease: an in vitro study. *Journal of Virology* **68**, 5765–5771.

Yamshchikov, V. F. & Compans, R. W. (1995). Formation of the flavivirus envelope: role of the viral NS2B-NS3 protease. *Journal of Virology* **69**, 1995–2003.

Yan, Y., Munshi, S., Li, Y., Sardana, V., Blue, J., Johns, B., Cole, J., Steinkuhler, C., Tomei, L., Francesco, R. D., Kuo, L. C. & Chen, Z. (1998). Three dimensional structure of BK strain hepatitis C virus (HCV) NS3-NS4A complex in a hexagonal crystal form at 2·2 Å resolution. Unpublished.

Zhang, R., Durkin, J., Windsor, W. T., McNemar, C., Ramanathan, L. & Le, H. V. (1997). Probing the substrate specificity of hepatitis C virus NS3 serine proteinase by using synthetic peptides. *Journal of Virology* **71**, 6208–6213.