Virus-Specific Cofactor Requirement and Chimeric Hepatitis C Virus/GB Virus B Nonstructural Protein 3

NANCY BUTKIEWICZ,1 NANHUA YAO,2 WEIDONG ZHONG,1 JACQUELYN WRIGHT-MINOGUE,1 PAUL INGRAVALLO,1 RUMIN ZHANG,2 JAMES DURKIN,2 DAVID N. STANDRING,1 BAHIGE M. BAROUDY,1 DAVID V. SANGAR,2 STANLEY M. LEMON,3 JOHNSON Y. N. LAU,4 AND ZHI HONG1*

Departments of Antiviral Therapy1 and Structural Chemistry,2 Schering-Plough Research Institute, Kenilworth, New Jersey 07033-0539, and Department of Microbiology and Immunology, The University of Texas Medical Branch at Galveston, Galveston, Texas 77555-10193

Received 22 November 1999/Accepted 22 January 2000

GB virus B (GBV-B) is closely related to hepatitis C virus (HCV) and causes acute hepatitis in tamarins (Saguinus species), making it an attractive surrogate virus for in vivo testing of anti-HCV inhibitors in a small monkey model. It has been reported that the nonstructural protein 3 (NS3) serine protease of GBV-B shares similar substrate specificity with its counterpart in HCV. Authentic proteolytic processing of the HCV polyprotein junctions (NS4A/4B, NS4B/5A, and NS5A/5B) can be accomplished by the GBV-B NS3 protease in an HCV NS4A cofactor-independent fashion. We further characterized the protease activity of a full-length GBV-B NS3 protein and its cofactor requirement using in vitro-translated GBV-B substrates. Cleavages at the NS4A/4B and NS5A/5B junctions were readily detectable only in the presence of a cofactor peptide derived from the central region of GBV-B NS4A. Interestingly, the GBV-B substrates could also be cleaved by the HCV NS3 protease in an HCV NS4A cofactor-dependent manner, supporting the notion that HCV and GBV-B share similar NS3 protease specificity while retaining a virus-specific cofactor requirement. This finding of a strict virus-specific cofactor requirement is consistent with the lack of sequence homology in the NS4A cofactor regions of HCV and GBV-B. The minimum cofactor region that supported GBV-B protease activity was mapped to a central region of GBV-B NS4A (between amino acids Phe22 and Val36) which overlapped with the cofactor region of HCV. Alanine substitution analysis demonstrated that two amino acids, Val27 and Trp31, were essential for the cofactor activity, a finding reminiscent of the two critical residues in the HCV NS4A cofactor, Ile25 and Ile29. A model for the GBV-B NS3 protease domain and NS4A cofactor complex revealed that GBV-B might have developed a similar structural strategy in the activation and regulation of its NS3 protease activity. Finally, a chimeric HCV/GBV-B bifunctional NS3, consisting of an N-terminal HCV protease domain and a C-terminal GBV-B RNA helicase domain, was engineered. Both enzymatic activities were retained by the chimeric protein, which could lead to the development of a chimeric GBV-B virus that depends on HCV protease function.

GB virus B (GBV-B) is a single-stranded (ss) positive-sense RNA virus associated with GB agent hepatitis (47, 60). It infects tamarins (Saguinus species) and causes acute hepatitis in naive animals (13). Phylogenetic tree analysis and the genome organization of GBV-B suggest that this virus belongs to the Flaviviridae family which at present consists of three genera: Flavivirus, Pestivirus, and Hepacivirus (47, 60). Like other members of this family (27, 52), the genome of GBV-B encodes a single large polyprotein of approximately 2,860 amino acids. The polyprotein is likely processed into several structural (C, E1, E2, and p7) and nonstructural (NS2, NS3, NS4A, NS4B, NS5A, and NS5B) proteins by either host- or virus-encoded proteases (47). Among all animal viruses, GBV-B shares closest nucleotide homology with hepatitis C virus (HCV) (47, 60). Like HCV, GBV-B is hepatotropic and causes liver disease in susceptible primates, which makes it a candidate as a separate member in the genus Hepacivirus according to a recent suggestion by the International Committee on Viral Taxonomy (54). In addition, GBV-B shares a similar IRES structure and function (53), as well as polyprotein organization (47, 48), with HCV.

HCV is the leading etiological agent of non-A non-B hepatitis (1, 12). More than 170 million people worldwide are infected by HCV. About 80% of patients with acute HCV infection will progress to chronic hepatitis; 20% of these will develop cirrhosis, and 1 to 5% of these will develop hepatocellular carcinoma (67). A recently completed population-based survey revealed that in the United States alone, the overall prevalence of anti-HCV was 1.8%, which corresponds to an estimated 3.9 million individuals infected by HCV nationwide. More than 74% of these seropositive individuals test positive for serum HCV RNA, indicating that at least 2.7 million persons are chronically infected (2). Current therapies with alpha interferon (IFN-α) alone and the combination of IFN-α and Ribavirin (Rebetron; Schering-Plough Corp., Kenilworth, N.J.) have been shown to be effective in no more than 41% of the patients with chronic HCV infection (44, 51). Vaccine development has been hampered by a high degree of antigenic variation and the lack of protection against reinfec-

* Corresponding author. Mailing address: Antiviral Therapy, K-15-4/4945, Schering-Plough Research Institute, 205 Galloping Hill Rd., Kenilworth, NJ 07033-0539, Phone: (908) 740-3152. Fax: (908) 740-3918. E-mail: zhi.hong@spcorp.com.
protease-RNA helicase, as well as NS5B polymerase, have been
grown to delineate important features relevant to the develop-
ment of potent and specific anti-HCV inhibitors (32, 33, 38, 42,
43, 71, 74).

The recent development of infectious molecular clones of HCV (4, 25, 34, 72, 73) has led to the establishment of a cell-based RNA replication system using an antibiotic-selected subgenomic RNA replicon (41). This may prove to be valuable for the in vitro evaluation of inhibitors targeting HCV repli-

cation. Unfortunately, the search for better and more specific anti-

tiral agents will be complicated by the lack of a readily

available animal model for hepatitis C other than the chim-

panzee, which is associated with limited availability and ex-

tremely high cost. This poses an almost insurmountable obsta-

cle for in vivo testing of anti-HCV inhibitors before clinical

trials in humans. The need for a small surrogate animal model

is urgent. The GBV-B–tamarin model may provide an effective

in vivo system reminiscent of the use of the woodchuck hepato-

titis model, a strategy that has been successfully developed for

anti-HBV drug development.

However, GBV-B is relatively new and uncharacterized. Limited studies of the GBV-B NS3 serine protease (57), RNA helicase (77), and NS5B RNA-dependent RNA polymerase (W. Zhong and Z. Hong, unpublished results) have revealed that there are many structural and enzymatic properties shared by GBV-B and HCV, further supporting the notion of GBV-B as a surrogate virus for HCV. The report by Scarselli et al.,
demonstrated that GBV-B NS3 protease shared similar sub-

strate specificity with that of HCV and cleaved the HCV polyprotein at the correct junction sites in an HCV NS4A-

dependent fashion (57). Although it has not been shown that the HCV NS3 protease is capable of reciprocal processing of the GBV-B polyprotein junctions, this finding provides encour-

gagement for developing the GBV-B–tamarin model and for

the in vivo testing of HCV protease inhibitors.

In this report, we address five key questions related to a recombinant GBV-B NS3 protein produced in E. coli. (i) Does the GBV-B NS3 protease process the GBV-B polyprotein as predicted by Scarselli et al. based on the use of the heterolo-

gous HCV polyprotein substrates (56)? (ii) Does HCV NS3 protease cleave the GBV-B polyprotein at the predicted junc-


tion sites? (iii) Is there an NS4A cofactor requirement by the

GBV-B NS3 protease? (iv) If so, what is the minimal region of

the NS4A that supports the GBV-B cofactor activity. (v) What are the critical amino acids essential for the cofactor activity?

Our results support the notion that HCV and GBV-B share similar NS3 protease specificities, while retaining a virus-spe-


cific NS4A cofactor requirement. A model for the GBV-B NS3

protease domain and NS4A cofactor complex is proposed that

shows similar structural features in the activation and sta-

bilization of the protease domain. Finally, a chimeric HCV/

GBV-B bifunctional NS3, consisting of an N-terminal HCV

protease domain and a C-terminal GBV-B RNA helicase do-

main, was engineered to explore the possibility of developing a

chimeric GBV-B virus that depends on HCV serine protease

function.

MATERIALS AND METHODS

Cloning of expression plasmids. The full-length GBV-B NS3 cDNA was iso-

lated directly from infected tamarin serum as described previously (76). Extrac-

tion of RNA from serum was by standard procedures. First, 0.2 μl of serum (GB

agent pool mystar 666, 8/93, kindly supplied by J. Bukh, National Institutes of

Health) was diluted with 100 μl of calf fetal serum and extracted using the Trizol

system (GIBCO BRL, Rockville, Md.). The pellet was dissolved in 10 mM

dithiothreitol (DTT) containing 20 U of RNasin (Promega, Madison, Wis.) per

ml. The selection of primers for cDNA synthesis and PCR amplification was

performed using the Superscript reverse transcriptase (GIBCO BRL) and the Advantage cDNA polymerase mix (Clontech, Palo Alto, Calif.). Four subgenomic cDNA

fragments were amplified covering the entire published genome. DNA fragments containing the GBV-B cleavage sites were isolated by PCR from cDNA clones of

GBV-B. The primers used to amplify the NS4A/4B junction were 5'-ATATGG

ATCCGGGTGCTACGGCCAGTGTG-3' and 5'-ATATAAAGCTTCACT

GGACCCATTTTGCGCCTCC-3'. The resulting PCR fragment was directly

cloned into pET-28a (Novagen, Madison, Wis.) between the

BamHI and HindIII sites. The primers for the GBV-B NS4B/5A cleavage site were 5'-AAATGGCC

TAGCGGTGACGTGGCCACCTAGGA3' and 5'-ATATGGATCAT

CACACCAGGTGTGTG-3'. The PCR amplified fragment was cloned into pET/

NS5B/CT-His (17) by replacing the NS5B coding region with that of the

NS4B/5A substrate between the

NheI and BglII sites.

The GBV-B NS3 protease domain and NS4A/4B, NS4B/5A, and NS5A/5B, respectively, were cloned as described by

Butkiewicz et al. (10). A chimeric HCV/GBV-B NS3 gene was constructed by

joining two PCR fragments: one consisted of the HCV-1a (H) NS3 protease
coding region (amino acids [aa] 1 to 190); the other consisted of the GBV-B

RNA helicase coding region (aa 190 to 620). The primers used to amplify the NS4A/4B junction were 5'- ATATAAAGCTTCACT

GGACCCATTTTGCGCCTCC-3'. The resulting PCR fragment was directly

cloned into pET-28a (Novagen, Madison, Wis.) between the

BamHI and HindIII sites. The full-length GBV-B NS3 was isolated by PCR from

cDNA clones of

GBV-B. The primers used to amplify the NS4A/4B junction were 5'-ATATGG

ATCCGGGTGCTACGGCCAGTGTG-3' and 5'-ATATAAAGCTTCACT

GGACCCATTTTGCGCCTCC-3'. The resulting PCR fragment was directly

cloned into pET-28a (Novagen, Madison, Wis.) between the

BamHI and HindIII sites. The primers for the GBV-B NS4B/5A cleavage site were 5'-AAATGGCC

TAGCGGTGACGTGGCCACCTAGGA3' and 5'-ATATGGATCAT

CACACCAGGTGTGTG-3'. The PCR amplified fragment was cloned into pET/

NS5B/CT-His (17) by replacing the NS5B coding region with that of the

NS4B/5A substrate between the

NheI and BglII sites. The GBV-B NS3A/5B was cloned similarly as that of NS4B/5A with the primers 5'-AAATGCGTAG

CCACCTAATTTGGCTGATGCAC-3' and 5'-ATATGGATCATCCATCTTT

CAACACACCATCATTTCACTC-3'. The GBV-B NS3B/CT-His coding region shown in Fig. 1 were numbered according to amino acid positions published previously (47, 60). Plas-
mids pBVB1, pBdB03, and pTS102 encoding the HCV NS3 cleavage sites

NS4A/4B, NS4B/5A, and NS5A/5B, respectively, were cloned as described by

Butkiewicz et al. (10). A chimeric GBV/B-HCV NS3 gene was constructed by

joining two PCR fragments: one consisted of the HCV-1a (H) NS3 protease
coding region (amino acids [aa] 1 to 190); the other consisted of the GBV-B

RNA helicase coding region (aa 190 to 620). The resulting chimeric cDNA was

cloned into pET/NS3-His-By replacing the full-length GBV-B NS3 between the

NheI and BamHI sites. The full-length HCV NS3 was isolated by PCR from

pHRM/HCY (1-301) (kindly provided by Charles Rice, Washington Univer-

sity, St. Louis, Mo.) and cloned into the BamHI site in pQE40 (Qiagen, Valencia,

Calif.). A His tag as well as an enterokinase cleavage site (MRGSHHHHHMS

DYKDDDDK) was inserted at the beginning of the NS3 gene to facilitate

purification and removal of the His tag. All plasmids were sequenced, and their

coding regions (shown in Fig. 1) were verified with an ABI Prism 377 DNA

sequencer with XL upgraded gels (PE Applied Biosystems, Foster City, Calif.).

FIG. 1. Schematic illustration of various plasmid constructs and cofactor peptides used in this study. The numbers above or below each construct or
deep indicate the amino acid positions corresponding to the published full-

length polyprotein for each virus. The predicted molecular masses for each

substrate and its cleavage products were indicated under each substrate plasmid.
Expression and purification. The full-length GBV-B NS3, HCV/GBV-B chimeric NS3, and HCV NS3 DNAs were freshly transformed into appropriate strains [E. coli JM109(DE3) for GBV-B NS3 and HCV/GBV-B NS3; E. coli M15(pREP4) for HCV NS3]. Protein production was induced at 24°C for 4 h by 0.2 mM isopropyl-β-D-thiogalactoside (IPTG) in Luria-Bertani media. Each protein (GBV-B NS3-His, HCV/GBV-B NS3-His or HCV H53-S3) was purified similarly on a nickel-chelated (Ni-nitriloacetic acid) affinity column, followed by a gel filtration column (Superdex 200; Amersham Pharmacia Biotech). The protein was stored in small aliquots at −70°C in the presence of 10% glycerol and 5 mM DTT.

In vitro translation of GBV-B and HCV substrates. Plasmids pPNBnae, pBII003, and pS7102, encoding the HCV NS4A/4B, NS4B/5A, and NS5A/5B cleavage sites, respectively, were linearized as described by Butkiewicz et al. (10). Plasmids containing the GBV-B NS4A/4B, NS4B/5A, and NS5A/5B cleavage sites were all linearized with XhoI. All plasmids were transcribed in vitro with T7 RNA polymerase. The in vitro-transcribed RNAs were translated in rabbit reticulocyte lysates (Promega) in the presence of [35S]methionine (Amersham Pharmacia Biotech) at 30°C for 1 h according to the supplier’s recommendation. All translation reactions were terminated by adding DNA-free RNase (Boehringer Ingelheim, Indianapolis, Ind.) and cycloheximide for 15 min at 30°C.

trans-cleavage translation assays. Standard translation assays for GBV-B NS3 protease were performed in a 20-μl reaction volume, initiated by adding 2 μl of [35S]-labeled translated substrate to purified protease in a final reaction mixture containing 50 mM Tris-Cl, pH 7.5; 50 mM NaCl; 0.1% lauryl maltoside; 10% glycerol; and 1 mM DTT. HCV protease was assayed similarly in a final reaction volume containing 50 mM MOPS, pH 7.5; 300 mM NaCl; 0.1% NP-40; 10% glycerol; and 1 mM DTT. Reactions were incubated as indicated in Fig. 3 (lanes 1 to 2 h at 30°C) and then terminated by adding an equal volume of 2× Laemmli sample buffer followed by boiling for 3 min. Cleavage products were separated by sodium dodecyl sulfate (SDS)–15% polyacrylamide gel electrophoresis (PAGE), detected by a PhosphorImager, and quantified by the ImageQuant software (both from Molecular Dynamics, Sunnyvale, Calif.).

Preparation of substrates for RNA helicase assay. RNA helicase substrates containing two complementary RNA strands were annealed and gel purified. Both strands were separately transcribed using the bacterichorionic T7 RNA polymerase (Promega) according to the manufacturer’s instructions. The annealed double-stranded RNA (dsRNA) substrates were end labeled using [γ-32P]ATP and T4 polynucleotide kinase (Amersham Pharmacia Biotech). The standard substrate (5′-3′) was prepared as follows. Plasmid digested plasmid pGEM-1 was transcribed with SP6 RNA polymerase to generate a 98-base RNA strand. XbaI-digested plasmid pSp65 was transcribed with SP6 RNA polymerase to generate a 38-base strand. The two RNA strands, containing a 29-base complementary region, were annealed and purified according to protocols published previously by other groups (31, 37, 65).

RNA helicase assay. The standard RNA helicase assay was carried out in a 20-μl reaction volume containing various concentrations of NS3 protein as indicated. A total of 50 fmol of labeled dsRNA substrates was added to the reaction buffer containing 50 mM Tris-HCl (pH 7.1), 1 mM MgCl2, 2 mM ATP, 2 mM DTT, 0.1 mg of bovine serum albumin per ml, and 4 U of RNasin RNase inhibitor (Promega). The reaction mixture was incubated at 37°C for 1 h, and the reaction was terminated by heating the mixture of 5 μl of the reaction buffer (100 mM HEPES, pH 7.5; 20 mM EDTA; 0.1% NP-40; 30% glycerol; 0.3% bromophenol blue). The RNA products were electrophoresed on a 4 to 20% polyacrylamide-TBE gel (Novex, San Diego, Calif.). The gel was dried, and 32P-labeled release strands were detected by autoradiography and quantified with a PhosphorImager.

RESULTS

A previous report (56) described the protease activity of GBV-B NS3 in which a catalytic domain was produced from E. coli. The study concluded that the GBV-B NS3 protease was able to recognize the authentic cleavage sites in the HCV polyprotein and to process the HCV polyprotein at the correct junctions. This finding was supported by a computer model prediction that the GBV-B protease possesses a similar substrate recognition site (the S1 pocket) for a small P1 residue (Cys) next to the scissile bond (49). The amino acid that defines the S1 pocket specificity (Phe239) is conserved between the HCV and GBV-B NS3 molecules (57). However, the reported results were surprising in that the GBV-B NS3 protease did not seem to require a cofactor for activity, a hallmark of flavivirus-like viruses (6, 14, 15, 39, 63, 64, 66). This lack of cofactor dependency could be due to the fact that only a truncated catalytic domain was used in the characterization of the protease activity which might be different from that of the full-length NS3 as demonstrated with HCV (20). Hence, we elected to further characterize the GBV-B protease activity using a full-length GBV-B NS3 molecule produced in E. coli as previously described (76).

Cleavage of HCV polyprotein junctions by the full-length GBV-B NS3 depends on the presence of a cofactor. We initially studied the role of GBV-B NS4A in GBV-B-mediated trans cleavages by using in vitro-translated polyprotein substrates containing the NS4A/4B, NS4B/5A, and NS5A/5B cleavage sites of the HCV polyprotein. We arbitrarily chose a 31-aa peptide (aa 16 to 46) from the central region of GBV-B NS4A (Δ4A31) to test whether the protease activity of the full-length GBV-B NS3 required an NS4A cofactor for optimal activity, as is the case with HCV. The in vitro-translated substrates were incubated as described previously (10) with the full-length GBV-B NS3 protease in the presence or absence of the Δ4A31 peptide. The results shown in Fig. 2 demonstrated that there was little processing of the HCV substrates at the NS4A/4B, NS4B/5A, and NS5A/5B junction sites in the absence of the NS4A peptide. As a control for appropriate polyprotein processing in this experiment, the full-length HCV NS3/4A/4D was described by Sali et al. (55) was used to produce the similar cleavage products as size markers (Fig. 2, lanes 2). However, processing of these substrates was increased and readily detectable with the addition of the NS4A peptide (compare Fig. 2, lanes 4 and 5), suggesting that, like HCV, GBV-B has also acquired a cofactor function in the central region of NS4A.

GBV-B NS3 protease processes its own polyprotein at the predicted sites. Several plasmids encoding the predicted GBV-B substrate cleavage sites at NS4A/4B, NS4B/5A, and NS5A/5B were constructed (Fig. 1) in order to evaluate homologous processing by GBV-B NS3. Details of the cloning and sequences of these substrate-encoding plasmids are given in Materials and Methods. In vitro trans cleavage assays were established with these substrates similarly to those described previously (10). In these experiments, a mutant GBV-B NS3 protease in which the predicted active site serine 137 was replaced by an alanine (S137A) served as a negative control for the protease activity. The results are shown in Fig. 3. Cleavage at the NS4A/NS4B junction by the wild-type GBV-B NS3 was readily detectable only in the presence of NS4A cofactor (Fig. 3A, lane 3), producing an NS4B product with a predicted molecular mass of 10 kDa. Similarly, processing at the NS5A/NS5B junction was also NS4A cofactor dependent, producing appropriately sized NS5A and NS5B products of 22 and 20 kDa, respectively (Fig. 3C). As expected, no processing was observed for the NS5A317A mutant protease with either substrate (lanes 4 in Fig. 3A and C). For an unknown reason, the wild-type GBV-B NS3 did not cleave the NS4B/5A substrate, either with or without the NS4A cofactor (Fig. 3B, lanes 2 and 3). The full-length NS3 of GBV-B consisted of a consensus sequence when compared to 12 independently isolated NS3 cDNA clones. It is also identical to the published sequences (47, 60) as well as to that of the infectious clone (9). In the report by Scarselli et al. the cDNA coding for the catalytic domain of NS3 was synthesized based on the published sequences. Thus, our full-length GBV NS3 consists of a protease domain that is identical to the one characterized by Scarselli et al. The only difference is the presence of the RNA helicase domain and a C-terminal His tag in our construct. The lack of activity was not due to the presence of the His tag, because either replacing the His tag at the N terminus of NS3 or removal of it by thrombin cleavage failed to improve the protease activity (data not shown). Alternatively, the in vitro-translated GBV-B NS4B/5A substrate may be somehow defec-
tive in a way that prevents proper cleavage by the protease. Further investigation of this issue is under way. Given the poor cleavage efficiency at this site in HCV, requiring 100- to 1,000-fold more enzyme than that required to cleave the NS4A/4B and NS5A/5B junctions (10, 61), it is not surprising that the cleavage at GBV-B NS4B/5A may be too inefficient to be detectable.

**NS4A cofactor-dependent cleavage of GBV-B substrates by HCV NS3 protease.** Having shown the NS4A cofactor-dependent GBV-B NS3 protease activity on HCV, as well as the GBV-B substrates, it was important to determine whether the HCV NS3 protease could reciprocally process the heterologous GBV-B polyprotein cleavage sites. In these studies, a full-length HCV NS3 was expressed and purified to 95% homogeneity from *E. coli*. Its activity for proteolytic processing was tested in the presence or absence of a 13-aa cofactor peptide from HCV NS4A (aa 22 to 34, Δ4A13) which has been shown to be sufficient for cofactor activation of this NS3 protease (10, 39, 58, 64). The results (Fig. 4A) showed a dose-dependent and NS4A cofactor-enhanced cleavage activity at the NS4A/4B junction site (lanes 6 to 9). In the absence of NS4A cofactor, some processing occurred with the highest concentration of HCV protease tested at 2.5 μM (Fig. 4A, lane 2). Cleavage of the GBV-B NS5A/5B site was inefficient and required 10 μM of HCV NS3 protease as well as the HCV NS4A cofactor (Fig. 4B, lanes 2 and 3). Again, no processing was observed at the predicted NS4B/5A site (data not shown).

To map the exact cleavage site, several synthetic peptides containing the predicted cleavage sites were generated. High-pressure liquid chromatography and mass spectrometric analyses of the cleavage products confirmed the prediction of Cys as the

---

**FIG. 2.** Cofactor-dependent trans cleavage activity of GBV-B NS3 protease on HCV substrates. 35S-labeled in vitro-translated substrates from HCV cleavage sites were translated as described in Materials and Methods: NS4A/NS4B (Δ4A/Δ4B) (A), NS4B/NS5A (Δ4B/Δ5A) (B), and NS5A/NS5B (Δ5A/Δ5B) (C). Labeled substrates and purified GBV-B protease were incubated in the presence or absence of 20 μM GBV-B NS4A peptide (aa 16 to 46, Δ4A13) for 1.5 h at 30°C. Final concentrations of GBV-B protease were 250 nM for NS4A/4B and 5 μM for NS4B/5A and NS5A/5B. In parallel, cleavages of HCV substrates by purified HCV NS3/4A are shown in each panel (lane 2). Samples were analyzed on an SDS-15% PAGE gel and detected by phosphorimaging scan.

**FIG. 3.** Cofactor-dependent trans cleavage activity of GBV-B NS3 protease on GBV-B substrates. 35S-labeled in vitro-translated substrates from GBV-B cleavage sites were mixed with 2 μM purified protease in the presence or absence of 20 μM GBV-B NS4A peptide (Δ4A13): Δ4A/Δ4B (A), Δ4B/Δ5A (B), and Δ5A/Δ5B (C). A mutant (NS3S137A) GBV-B protease was also tested (lane 4 of each panel). Proteins were separated and analyzed as described in the legend to Fig. 2.
P1 residue at the NS4A/4B junction (data not shown) as described previously (57).

**Mapping the minimal cofactor domain within GBV-B NS4A.**

The results described above demonstrated that the GBV-B cofactor activity resides in the central region of NS4A from aa 16 to 46. To determine the minimum cofactor region required by GBV-B NS protease, we synthesized a series of truncated peptides derived from this central region of GBV-B NS4A. These synthetic peptides were tested for their ability to activate GBV-B NS3 protease activity at the NS4A/4B junction site. As summarized in Fig. 5, the C-terminal half (an 18-mer peptide corresponding to aa 29 to 46) of the central 31-aa peptide did not possess detectable cofactor activity since it failed to enhance the protease activity. On the other hand, a 17-mer peptide representing aa 20 to 36 of NS4A retained cofactor activity. Further progressive truncations from the N terminus of this peptide produced a 15-mer peptide (aa 22 to 36), which represented the minimum region that could efficiently activate the protease in these trans cleavage assays. Removal of additional residues from either terminus severely reduced this cofactor activity. It is interesting to note that this minimum cofactor region in GBV-B NS4A (aa 22 to 36) overlaps with that of

![Diagram](https://via.placeholder.com/150)

**FIG. 4.** Cofactor-dependent trans cleavage activity of HCV NS3 protease on GBV-B substrates. (A) trans cleavage reaction using GBV-B NS4A/4B as the substrate. Purified HCV full-length NS3 protease was mixed with translated substrate in the presence or absence of 100 μM HCV NS4A cofactor peptide (Δ4AΔ4B). Decreasing concentrations of protease were tested in this assay (indicated by a triangular descending slope: 2.5, 0.25, 0.025, and 0.0025 μM). (B) trans cleavage reaction using GBV-B NS5A/5B as the substrate. HCV NS3 protease at 10 μM was tested in the presence or absence of 100 μM NS4A cofactor peptide (Δ4AΔ4B).

**FIG. 5.** Mapping the minimum region within GBV-B NS4A required for the cofactor activity. Sequences of the entire HCV and GBV-B NS4A proteins are aligned at the top with the minimum cofactor region for HCV NS4A underlined. Vertical lines represent identity while colons represent similarity between the two viruses within the NS4A region. Dashes represent a gap created to maximize the homology alignment. To compare the cofactor activity, 35S-labeled GBV-B NS4A/4B substrate (Δ4AΔ4B) was incubated with 1.5 μM GBV-B protease in the presence of various GBV-B NS4A peptides as shown in the figure. Cleavage products were then analyzed and detected by phosphorimaging and quantified by ImageQuant software (Molecular Dynamics). The cofactor activity was graded by comparing the enhancement of the protease activity to that of the background (cleavage activity in the absence of NS4A cofactor and scored as follows: ++++, 11- to 20-fold; ++, 5- to 10-fold; +, 2.5- to 5-fold; ±, 1.5- to 2-fold; and –, <1.5-fold of the background activity. Amino acids critical for cofactor activation of the NS3 protease are shown in boldface, larger type; the minimum GBV-B NS4A cofactor region is also underlined.

<table>
<thead>
<tr>
<th>Cofactor</th>
<th>Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>HCV</td>
<td>++</td>
</tr>
<tr>
<td>GBV-B</td>
<td>++</td>
</tr>
<tr>
<td>16-46</td>
<td>K-YLAIDTFGATCVRRCWITSVFGATVAPVV</td>
</tr>
<tr>
<td>29-46</td>
<td>RCWITSVFGATVAPVV</td>
</tr>
<tr>
<td>20-36</td>
<td>DTFGATCVRRCWITSV</td>
</tr>
<tr>
<td>21-36</td>
<td>TFGATCVRRCWITSV</td>
</tr>
<tr>
<td>22-33</td>
<td>FGATCVRRCWITSV</td>
</tr>
<tr>
<td>23-34</td>
<td>GATCVRRCWITSV</td>
</tr>
<tr>
<td>20-35</td>
<td>DTFGATCVRRCWITS</td>
</tr>
<tr>
<td>20-34</td>
<td>DTFGATCVRRCWIT</td>
</tr>
<tr>
<td>20-33</td>
<td>DTFGATCVRRCWIT</td>
</tr>
<tr>
<td>22-34</td>
<td>FGATCVRRCWIT</td>
</tr>
<tr>
<td>23-34</td>
<td>GATCVRRCWIT</td>
</tr>
<tr>
<td>21-32</td>
<td>TPGATCVRRCWMS</td>
</tr>
</tbody>
</table>
HCV which has been mapped to the central region (aa 22 to 33) of HCV NS4A (10, 39, 58, 64) (Fig. 5).

**NS4A cofactor peptide activity is virus specific.** The experiments described above demonstrated that the full-length GBV-B NS3 protease has the ability to trans-cleave the homologous GBV-B substrates as well as the corresponding heterologous HCV substrates. Conversely, the HCV full-length NS3 protease can cleave the heterologous GBV-B substrates in addition to the HCV substrates. Since this cleavage activity was cofactor dependent for both proteases, it was of interest to determine whether or not the required cofactor activity was virus specific. For this purpose, the NS4A/4B substrates from both GBV-B and HCV were chosen to evaluate cofactor specificity. Each NS3 protease was tested with its homologous substrate or the heterologous GBV-B substrates in the absence of NS4A peptide or in the presence of either HCV NS4A (Δ4A4B) or GBV-B NS4A (Δ4A4B) at 100 μM. The assay was performed using the NS4A/4B substrates from either HCV (A) or GBV-B (B). HCV NS3 was tested at 50 nM for both substrates; GBV-B protease was tested at 250 nM for HCV substrate and at 2 μM for GBV-B substrate.

![Image](https://via.placeholder.com/150)

**Figure 6.** NS4A cofactor activity is virus specific. HCV or GBV-B NS3 protease was tested in the absence of NS4A peptide or in the presence of either HCV NS4A (Δ4A4B) or GBV-B NS4A (Δ4A4B) at 100 μM. The assay was performed using the NS4A/4B substrates from either HCV (A) or GBV-B (B). HCV NS3 was tested at 50 nM for both substrates; GBV-B protease was tested at 250 nM for HCV substrate and at 2 μM for GBV-B substrate.

**Probing essential amino acid requirements for GBV-B NS4A cofactor activity.** To investigate how GBV-B NS4A cofactor might interact with and enhance the NS3 protease activity, the HCV NS3 protease structure was used as a template for the modeling of GBV-B protease structure using the program QUANTA (Molecular Simulations, Inc., San Diego, Calif.). The primary sequences of the NS3 protease domain between HCV and GBV-B share approximately 30% identity, whereas the full-length NS3 proteins share 40% identity. The secondary structure elements of the GBV-B protease model were transferred from the existing HCV NS3 protease crystal structures (33, 43, 71). Side chain atom coordinates of HCV were adopted to model and position those in the corresponding GBV-B protease. Those unresolved side chains were built by using the program CHARMM (Molecular Simulations), and the newly modeled molecule was relaxed to yield a smooth backbone with reasonable packing by molecular dynamics.

![Image](https://via.placeholder.com/150)

**Figure 8B shows the NS3/4A interaction of HCV.** The two critical residues (I35 and I39 in magenta) which are buried in two well-defined hydrophobic pockets in the protease core are highlighted. The I39 side chain occupies the hydrophobic pocket located between the two b-barrel subdomains and may rigidify the relative position of the subdomains that are connected by a single flexible loop. I39 may also pivot the insertion of the N-terminal half of NS4A cofactor into the N-terminal substitution at W31, almost completely abolished the cofactor-dependent cleavage, inhibiting the product formation by 96% as determined by using a PhosphorImager (Fig. 7C). Similarly, replacing V27 with alanine was also detrimental to the cofactor activity, inhibiting the cleavage by 86% (Fig. 7C). These results reveal a strikingly similar pattern of NS4A activation of NS3 protease between the two viruses, in that V27 and W31 of GBV-B mirror the two critical amino acids (I25 and I29) defining the HCV cofactor activity (10, 58) (Fig. 5). It appears that the hydrophobicity of the critical cofactor amino acids, as well as the spacing between these two residues, is important for the activation of NS3 proteases. Other amino acid residues in the cofactor regions may play a role in defining the viral specificity of cofactor function.

**Computer model for GBV-B NS3/4A interaction.** To further understand how GBV-B cofactor peptide might interact with and enhance the NS3 protease activity, the HCV NS3 protease structure was used as a template for the modeling of GBV-B protease structure using the program QUANTA (Molecular Simulations, Inc., San Diego, Calif.). The primary sequences of the NS3 protease domain between HCV and GBV-B share approximately 30% identity, whereas the full-length NS3 proteins share 40% identity. The secondary structure elements of the GBV-B protease model were transferred from the existing HCV NS3 protease crystal structures (33, 43, 71). Side chain atom coordinates of HCV were adopted to model and position those in the corresponding GBV-B protease. Those unresolved side chains were built by using the program CHARMM (Molecular Simulations), and the newly modeled molecule was relaxed to yield a smooth backbone with reasonable packing by molecular dynamics.

![Image](https://via.placeholder.com/150)

**Figure 8B shows the NS3/4A interaction of HCV.** The two critical residues (I35 and I39 in magenta) which are buried in two well-defined hydrophobic pockets in the protease core are highlighted. The I39 side chain occupies the hydrophobic pocket located between the two b-barrel subdomains and may rigidify the relative position of the subdomains that are connected by a single flexible loop. I39 may also pivot the insertion of the N-terminal half of NS4A cofactor into the N-terminal substitution at W31, almost completely abolished the cofactor-dependent cleavage, inhibiting the product formation by 96% as determined by using a PhosphorImager (Fig. 7C). Similarly, replacing V27 with alanine was also detrimental to the cofactor activity, inhibiting the cleavage by 86% (Fig. 7C). These results reveal a strikingly similar pattern of NS4A activation of NS3 protease between the two viruses, in that V27 and W31 of GBV-B mirror the two critical amino acids (I25 and I29) defining the HCV cofactor activity (10, 58) (Fig. 5). It appears that the hydrophobicity of the critical cofactor amino acids, as well as the spacing between these two residues, is important for the activation of NS3 proteases. Other amino acid residues in the cofactor regions may play a role in defining the viral specificity of cofactor function.
domain of NS3 protease. This is consistent with the observation that I_{29} is the most critical determinant of the NS3/4A activation (10). The hydrophobic side chain of I_{29}, another key residue, packs against two adjacent β strands (A1 and D1), which may cause the D1-E1 loop to shift from its position in NS3 upon NS4A complex formation. However, other mechanisms are possible, such as the one proposed by Lowe et al. (43), who favored the rearrangement of salt links among charged amino acids as the cause of the D1-E1 loop shift.

The modeled GBV-B complex structure reveals two strikingly similar hydrophobic pockets for docking the NS4A cofactor peptide and stringent complementarity between the protease and its cofactor (Fig. 8A). This is consistent with the mutation analysis described above (Fig. 7) that demonstrated the two critical amino acid determinants within the cofactor. The W_{31} side chain (corresponding to that of I_{29} in HCV) occupies a somewhat larger hydrophobic pocket between the β-barrel subdomains than that present in HCV. The hydrophobic side chain of V_{27} (corresponding to that of I_{25} in HCV) packs similarly as described in the previous figures. The alanine-substituted residue and the corresponding amino acid position (number) in NS4A are shown above each lane. The modeled GBV-B complex structure reveals two strikingly similar hydrophobic pockets for docking the NS4A cofactor peptide and stringent complementarity between the protease and its cofactor (Fig. 8A).

FIG. 7. Cofactor activity of alanine-substituted GBV-B NS4A peptides (Δ4AΔ4B). (A) Sequence of alanine-substituted peptides used in this study, with the alanine residue shown in boldface. The A24 wild type [A24(WT)] is the same as the wild-type peptide. (B) Cofactor activity tested in the trans cleavage assay. A 200 μM concentration of each peptide was used in each reaction containing the GBV-B NS4A/4B substrate and 1.5 μM GBV-B NS3 protease. Cleavage products were analyzed similarly as described in the previous figures. The alanine-substituted residue and the corresponding amino acid position (number) in NS4A are shown above each lane. Vertical arrows identify the amino acids critical for cofactor activity. (C) Quantitative analysis of the cofactor activities (shown in panel B) expressed as the percent inhibition of wild-type cofactor activity.
NS3 structure. Both the protease and the RNA helicase activities were subsequently characterized. The results (Fig. 9A) demonstrated that the chimeric NS3 retained protease activity capable of processing both GBV-B and HCV substrates (containing the NS4A/4B cleavage site). This protease activity required only the HCV NS4A cofactor peptide (Δ4A13) (lane 5) and not the GBV-B NS4A cofactor peptide (Δ4A17) (lanes 7 and 8). As a comparison, the full-length HCV NS3 protease

FIG. 8. Surface representation of GBV-B and HCV protease complexed with the corresponding 4A cofactors (shown as sticks). (A) GBV-B NS3/4A model. (B) HCV NS3/4A model. The electrostatic potential distribution is color coded: blue, positive; red, negative; white, neutral. The N-terminal 11 aa were removed to expose the detail interaction between NS4A and NS3 protease. The two critical Trp31 and Val27 (shown as magenta sticks in contrast to the rest of 4A colored yellow) reveal their location and complementary with the corresponding hydrophobic pockets, which is strikingly similar to that of HCV NS3/4A complex structure in panel B. The two essential residues Ile29 and Ile25, in HCV NS4A are also colored in magenta.

FIG. 9. HCV protease and GBV-B RNA helicase activities of the chimeric NS3 protein. (A) Protease activity. The GBV-B NS4A/4B (Δ4A/4B) substrate was mixed with the chimeric HCV/GBV-B NS3 at 2 μM (lanes 3, 5, and 7) or at 0.07 μM (lanes 4, 6, and 8) in the absence NS4A cofactor or the presence of either HCV NS4A cofactor (Δ4A13) or GBV-B NS4A cofactor (Δ4A17). The final concentrations of the cofactor peptides were 100 μM. The cleavage products were analyzed as described previously. (B) RNA helicase activity. The RNA helicase activity of the chimeric NS3 was compared to that of the native GBV-B NS3. The Δ symbol indicates that the sample was boiled before loading onto the gel. Increasing amounts (2, 5, and 10 pmol) of each protein were tested for the dsRNA unwinding activity. The released ssRNAs were separated and analyzed as described in Materials and Methods. The ascending triangular slopes indicate the increasing concentration of the NS3 proteins. The “−” symbol indicates the no-enzyme controls.
was used to demonstrate the HCV NS4A cofactor-dependent cleavage products (lane 9). Comparison of the RNA helicase activities also revealed that the chimeric NS3 retained the RNA helicase activity similar to that of native GBV-B NS3 (Fig. 9B). Thus, the chimeric NS3 acquired the HCV protease activity and retained the native GBV-B RNA helicase activity.

**DISCUSSION**

Virally encoded proteases play an essential role in viral replication. Inhibitors of viral proteases have been developed to inhibit viral replication. Several human immunodeficiency virus (HIV) protease inhibitors that effectively suppress the viral production and slow disease progression are approved for the treatment of HIV infection in combination with other antiviral components (46, 50). Like the HIV protease, the HCV NS3 serine protease has been considered an important target for the development of anti-HCV inhibitors. Unfortunately, small and relatively inexpensive animal models that support HCV replication are not available at the present time. Efforts to establish transgenic mouse models or xenografted SCID-hu mouse models have not yielded any reliable or convincing animal models that are suitable for anti-HCV drug development (8, 36, 45). The only animal model permissive for HCV infection in vivo is the chimpanzee, which is prohibitively expensive for routine drug discovery efforts. The recent discovery of GBV-B, a hepatotropic virus closely related to HCV, provides hope for a surrogate virus to evaluate anti-HCV inhibitors in a much less expensive small animal model.

Several HCV NS3 protease-dependent chimeric viruses have been created, providing necessary cell-based antiviral assays to evaluate the efficacy of potential inhibitors against the HCV protease (11, 18, 19, 23). Although these chimeric viruses may alleviate the pressing need for in vitro testing of cell permeation and antiviral efficacy of candidate inhibitors of HCV protease, these artificial systems in the most part fail to duplicate the polyprotein processing events occurring naturally during HCV infection. In addition, proper animal models have not been developed for these viruses. A chimeric HCV/GBV-B virus would provide an alternative solution that better mimics an HCV infection in a small primate model. The primary findings in this report further strengthen the hypothesis that the GBV-B-tamarin model is biologically relevant to the study of HCV and may serve as a good surrogate system for HCV infection (57). In addition, our comparative characterization of GBV-B NS3 protease and identification of a virus-specific GBV-B NS4A cofactor provide insights for the future development of a viable chimeric GBV-B/HCV virus that can be used in the tamarin model. In this improved version of the GBV-B-tamarin model, a built-in HCV NS3 protease domain as well as an HCV NS4A cofactor region would replace the resident GBV-B protease and NS4A cofactor, making the chimeric virus dependent upon the respective HCV functions.

Guided by a predicted computer model of the full-length GBV-B NS3 molecule as well as the full-length HCV NS3 structure, precise protease domain swapping was accomplished. The resulting chimeric HCV/GBV-B NS3 retained both protease and RNA helicase activities comparable to the unaltered or native functionality. The mapping of the GBV-B NS4A cofactor region provides necessary information for splicing the HCV NS4A cofactor into the GBV-B genome in lieu of its own cofactor. A unique sequence of 265 bases was recently identified at the 3′ end of the GBV-B genome (56; D. V. Sangar and S. M. Lemon, unpublished results), reminiscent of the 3′X element found at the 3′ end of HCV genome (35). This will allow the construction of full-length molecular clones of GBV-B (9) as well as chimeric HCV/GBV-B viruses. Since the HCV and GBV-B proteases share similar substrate specificity and can cross-process each other's polyprotein, it is likely that an HCV NS3 protease-dependent GBV-B would be capable of proper proteolytic processing and producing mature viral proteins that would be competent for viral replication. The identification of the virus-specific NS4A cofactor requirement (Fig. 6) suggests that a concomitant chimeric engineering of the NS4A domain would be necessary to achieve efficient heterologous polyprotein processing by the chimeric viruses. Such chimeric viruses would be valuable for both in vitro and in vivo testing of anti-HCV inhibitors.

The complete mechanistic pathway of NS4A activation of the NS3 protease is not fully understood. The crystal structures of HCV NS3 protease and NS4A cofactor complexes (33, 43, 71) have shown that two major conformational changes upon NS4A complexation occur: one involves the packing and ordering of the N-terminal 30 aa; the other causes a shift of the D1-E1 loop in the N-terminal β-barrel subdomain (43, 71). These conformational changes help to confine the catalytic triad in the active site to a proper orientation optimal for catalysis. Here, we propose a general mechanism for the activation of NS3 protease based on the structural and mutational analyses as well as on comparative studies of other related viruses. The two critical hydrophobic amino acids separated by three noncritical amino acids (ΦxxΦ, where Φ represents a bulky hydrophobic amino acid and x represents any amino acid) form a protease activation motif. This is further substantiated by the presence of this motif in the NS4A cofactors of HCV, GB viruses, and pestiviruses, as well as in NS2B cofactor region of flaviviruses (Fig. 10) (6, 7, 66). Based on the comparison between the structures of NS3 alone and NS3/4A complex (33, 42), the hydrophobic pocket corresponding to the second Φ (I29 in HCV NS4A) remains unchanged before and after NS4A complexation, while the pocket for the first Φ (I15 in HCV NS4A) undergoes significant conformational change. We thus hypothesize that the second hydrophobic amino acid of the “ΦxxΦ” motif makes contact with the protease. It plays an essential role by occupying the hydrophobic pocket between the two β-barrel subdomains and rigidifying the relative position of the subdomains which are connected by a single loop. This is followed by the insertion of the N-terminal half of NS4A cofactor into the N-terminal domain of NS3 protease, allowing the first hydrophobic amino acid of the activation motif to be inserted into a second hydrophobic pocket adjacent to two beta-strands (A1 and D1). This insertion may directly cause the D1-E1 loop shift upon cofactor complexation and stiffen the alignment of the catalytic triad in the active site. Another mechanism was proposed by Love et al. (43), who
favored the rearrangement of salt links among charged amino acids as the cause of D1-E1 loop shift. However, this cannot explain the results from the mutational analysis (10, 58) and the apparent lack of sequence conservation of the charged amino acids (Fig. 5 and 10). A recent study on HGV/GBV-C also revealed a similar cofactor requirement which was mapped to the central region of NS4A (6). With the alternating hydrophobic pattern, the HGV NS4A cofactor enhanced the NS3 protease activity. Interestingly, the HGV NS4A has been shown to be a weak cofactor of HCV NS3 protease, highlighting the high degree of functional conservation in related viruses (68). Despite the strict virus-specific cofactor requirement, a general strategy becomes apparent in the activation of NS3 protease by its NS4A cofactor.

ACKNOWLEDGMENTS

We thank Gregory Reyes for support and Patricia Weber for encouragement and critical reading of the manuscript. We also thank Yanhui Liu for mass spectrometry analysis and Angela Skelton and Eric Ferrari for assistance in molecular cloning and protein purification. This work was supported in part by a grant from the National Institute of Allergy and Infectious Diseases (U19-AI40035).

REFERENCES

57. Wiskerchen, M., and M. S. Collett. 1991. Pestivirus gene expression: protein p80 of bovine viral diarrhoea virus is a serine protease involved in polypro-