Functional Characterization of cis and trans Activity of the Flavivirus NS2B-NS3 Protease*§

Received for publication, December 11, 2006, and in revised form, March 1, 2007 Published, JBC Papers in Press, March 2, 2007, DOI 10.1074/jbc.M611318200

Aloke K. Bera‡, Richard J. Kuhn†, and Janet L. Smith§§

From the ‡Department of Biological Sciences, Purdue University, West Lafayette, Indiana 47907 and §§Life Sciences Institute and Department of Biological Chemistry, University of Michigan, Ann Arbor, Michigan 48109

Flaviviruses are serious human pathogens for which treatments are generally lacking. The proteolytic maturation of the 375-kDa viral polyprotein is one target for antiviral development. The flavivirus serine protease consists of the N-terminal domain of the multifunctional nonstructural protein 3 (NS3) and an essential 40-residue cofactor (NS2B40) within viral protein NS2B. The NS2B-NS3 protease is responsible for all cytoplasmic cleavage events in viral polyprotein maturation. This study describes the first biochemical characterization of flavivirus protease activity using full-length NS3. Recombinant proteases were created by fusion of West Nile virus (WNV) NS2B40 to full-length WNV NS3. The protease catalyzed two autolytic cleavages. The NS2B/NS3 junction was cleaved before protein purification. A second site at Arg459 Gly460 within the C-terminal helicase region of NS3 was cleaved more slowly. Autolytic cleavage reactions also occurred in NS2B-NS3 recombinant proteins from yellow fever virus, dengue virus types 2 and 4, and Japanese encephalitis virus. cis and trans cleavages were distinguished using a noncleavable WNV protease variant and two types of substrates as follows: an inactive variant of recombinant WNV NS2B-NS3, and cyan and yellow fluorescent proteins fused by a dodecamer peptide encompassing a natural cleavage site. With these materials, the autolytic cleavages were found to be intramolecular only. Autolytic cleavage of the helicase site was insensitive to protein dilution, confirming that autolysis is intramolecular. Formation of an active protease was found to require neither cleavage of NS2B from NS3 nor a free NS3 N terminus. Evidence was also obtained for product inhibition of the protease by the cleaved C terminus of NS2B.

Most flaviviruses, including West Nile virus (WNV), 2 yellow fever virus (YFV), dengue viruses, and Japanese encephalitis virus (JEV), cause severe human diseases. The plus-sense RNA genome of flaviviruses is a single open reading frame encoding a polyprotein precursor of ~3400 amino acids, consisting of three structural proteins (C, prM, and E) and seven nonstructural replication proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5) (Fig. 1A) (1–3). Signal sequences direct the proteolytic maturation of the polyprotein precursor at the NS2A/NS2B, NS2B/NS3, NS3/NS4A, and NS4B/NS5 junctions (Fig. 1A), as well as at internal sites within C, NS2A, NS3, and NS4A (9–12). In general, the viral protease has specificity for two basic residues (Lys-Arg, Arg-Arg, Arg-Lys or occasionally Gln-Arg) at the canonical P2 and P1 positions immediately preceding the cleavage site, followed by a small amino acid (Gly, Ser, or Ala) at the P1’ position (Fig. 1B).

The activity of NS2B-NS3 protease has been studied in vitro using purified, recombinant protease domains of dengue virus type 2 (DV2) and WNV NS3. These studies employed fusions of the NS2B cofactor peptide to truncated forms of NS3 (NS3pro) that included only the N-terminal protease domain, one-third of the full-length NS3, and excluded the C-terminal helicase domain (13–16). Among several fusion peptides tested, Gly4–Ser–Gly4 was found to be the optimal linker for DV2 NS2B-NS3pro (17, 18), and several groups have used this Gly4–Ser–Gly4 linker peptide for a variety of biochemical and mutagenesis studies of protease specificity (19–22).

Crystal structures of DV2 and WNV protease domains have been reported (23, 24). Crystal structures have also been reported for the helicase domains of YFV and DV2 (25, 26). The protease belongs to the chymotrypsin family with a classic Ser-His-Asp catalytic triad. The catalytic triad is arranged identically in structures with (24) and without (23) the NS2B40 cofactor. The NS2B40 cofactor contributes to the binding site for the P2 residue of the substrate, based on the structure of WNV NS2B-NS3 protease domain in complex with a substrate-based inhibitor (24).

Here we report single polypeptide variants of NS2B-NS3 protease from five flaviviruses. To our knowledge, this is the first study addressing flavivirus protease activity using the NS2B40 cofactor.
Cis and Trans Activity of Flavivirus NS2B-NS3 Protease

A Structural proteins Non-structural proteins

<table>
<thead>
<tr>
<th>C</th>
<th>prM</th>
<th>E</th>
<th>N</th>
<th>NS1</th>
<th>NS2A</th>
<th>NS2B</th>
<th>NS3</th>
<th>NS4A</th>
<th>NS4B</th>
<th>NS5</th>
</tr>
</thead>
<tbody>
<tr>
<td>WNV</td>
<td>AINRBS5QHBRGKCIAGVM</td>
<td>YFV</td>
<td>KNQLSRR7FDHVL/FCFL</td>
<td>DV2</td>
<td>GMN1INRNRGSIHMLT</td>
<td>DV4</td>
<td>GMN1INGRSTITLCLLP</td>
<td>JEV</td>
<td>AVNKRGRQ03SGN1GSTML</td>
<td>NS4B-NS5</td>
</tr>
<tr>
<td>NS2A-2B</td>
<td>AGL1ACQDRFGK2WATEVMA</td>
<td>YFV</td>
<td>CAF1TRFGQK1PV1NA1AA</td>
<td>DV2</td>
<td>ETL1SRTGQFDW1NA1AA</td>
<td>DV4</td>
<td>YLM11N1GQ1WPSN1EG1N</td>
<td>JEV</td>
<td>AG1M1CQ11F3WATEF1SA</td>
<td>NS2A (internal)</td>
</tr>
<tr>
<td>NS2B-NS3</td>
<td>VWG1T1Q1TQ2B1GVLWD1T6G</td>
<td>YFV</td>
<td>AG11L1FVPAG0GAV01VD1P1</td>
<td>DV2</td>
<td>AWD1W1V0000V1G1VWDS1P1</td>
<td>DV4</td>
<td>TL1W1M0Q1G1S1G1W1L1D1</td>
<td>JEV</td>
<td>EG01W1M1T1K1T1V1W1G1T1G</td>
<td>NS3 (internal)</td>
</tr>
<tr>
<td>NS3-NS4A</td>
<td>IKA0KD1FAG0B1Q1G1L1E1V1L1G</td>
<td>YFV</td>
<td>L0Nd1T1F0AB1G1V0LS1V1L1S</td>
<td>DV2</td>
<td>L0Nd1T1F0AB1G1V0LS1V1L1S</td>
<td>DV4</td>
<td>L0Nd1T1F0AB1G1V0LS1V1L1S</td>
<td>JEV</td>
<td>EG01W1M1T1K1T1V1W1G1T1G</td>
<td>NS4A (internal)</td>
</tr>
</tbody>
</table>

Inactive proteases with alanine substitutions at Ser135 (WNV, DV2, DV4, and JEV) or Ser138 (YFV) in the catalytic triad of NS3 were made by site-directed mutagenesis. For WNV, the substitution was generated by PCR from pWNV-NS2B40-G4SG4-HM-NS3FL and primers WNV-S135A-F and WNV-S135A-R, yielding pWNV-NS2B40-G4SG4-HM-NS3FL/S135A. The PCR product was digested with DpnI before transformation.

The plasmid pCYFP28 encoding cyan (CFP) and yellow fluorescent proteins (YFP) separated by multiple restriction sites was provided by Todd W. Geders (from our laboratory), in the pET28 vector with a C-terminal His tag. The annealed product of S-WNV-2B/3–1 and S-WNV-2B/3–2 (supplemental Table 1) encoding the sequence LQYTKR/GGVLWD1 between BamHI and HindIII restriction sites was ligated into pCYFP28 to generate pWNV-CFP-LQYTKR/GGVLWD1-YFP (or simply pWNV-CFP-2B/3-YFP). Constructs encoding other substrates were made similarly using primers listed in supplemental Table 1. The composition of all constructs was verified by DNA sequencing.

MATERIALS AND METHODS

Plasmid Construction—Constructs encoding WNV (strain NY99) proteins were derived from pWN-CG (27), a kind gift of Richard Kinney, Centers for Disease Control and Prevention. For the fusion of the C-terminal 79 residues (residues 53–131) of NS2B with full-length NS3, a 2154-nucleotide fragment was amplified from pWN-CG, using primers WNV-NS2B53-F and WNV-NS3FL619STOP-R (supplemental Table 1), and cloned between NdeI and BamHI restriction sites in pET28 (Stratagene) to generate pNS2B 53–131-NS3FL (simply pWNV-NS2B79-NS3FL). Other WNV plasmids were constructed in two steps. First, an 1857-nucleotide fragment was amplified from pWN-CG by primers WNV-NS31-F and WNV-NS3619STOP-R and cloned into pET28 between NdeI and BamHI restriction sites to generate pWNV-NS3FL encoding full-length NS3. Then fragments encoding the 41-residue protease cofactor within NS2B (residues 53–93) were amplified from pWN-CG using primers WN-NS2B53-F and either WNV-NS2B93-R or WNV-NS2B95-G4SG4-R. The fragments were digested with NdeI and cloned into NdeI-restricted pWNV-NS3FL, to generate either pWNV-NS2B53–93-HM-NS3FL (simply pWNV-NS2B40-HM-NS3FL) or pWNV-NS2B53–95-G4SG4-HM-NS3FL (simply pWNV-NS2B40-G4SG4-HM-NS3FL).

Plasmids encoding the active proteases for YFV (pYFV-NS2B40-G4SG4-HM-NS3FL), DV2 (pDV2-NS2B40-G4SG4-ASR-NS3FL), dengue virus type 4 (DV4, pDV4-NS2B40-G4SG4-ASR-NS3FL), and JEV (pJEV-NS2B40-G4SG4-ASR-NS3FL) were constructed in two steps, as described for the WNV protease. PCR primers are listed in supplemental Table 1. For the YFV, and DV4 constructs, an Ndel site between the sequences for the NS2B cofactor and NS3 encoded the amino acids His-Met. For the DV2 and JEV constructs, an NheI site was used, corresponding to Alase and followed by a basic residue “R,” corresponding to the C-terminal residue of NS2B. The plasmid for YFV, pACNR/FLYF (28), was provided by Charles Rice, Laboratory of Virology and Infectious Disease, The Rockefeller University, and plasmids for DV2 and DV4 were from Richard Kinney, and a partial cDNA clone of JEV was from Tsutomu Takegami, Medical Research Institute of Kanazawa Medical University.

Gene Expression and Protein Purification—The plasmid pCYFP28 encoding cyan (CFP) and yellow fluorescent proteins (YFP) separated by multiple restriction sites was provided by Todd W. Geders (from our laboratory), in the pET28 vector with a C-terminal His tag. The annealed product of S-WNV-2B/3–1 and S-WNV-2B/3–2 (supplemental Table 1) encoding the sequence LQYTKR/GGVLWD1 between BamHI and HindIII restriction sites was ligated into pCYFP28 to generate pWNV-CFP-LQYTKR/GGVLWD1-YFP (or simply pWNV-CFP-2B/3-YFP). Constructs encoding other substrates were made similarly using primers listed in supplemental Table 1. The composition of all constructs was verified by DNA sequencing.

FIGURE 1. Flavivirus protease cleavage sites. A, schematic diagram of the polyprotein with cleavage sites for the viral protease indicated by solid arrowheads. B, sequences of cleavage sites for flavivirus NS2B-NS3 proteases. Cleavage by the viral protease occurs after the boldface underlined residue in the “P1” position, in most cases R.
the recombinant protein was induced by addition of isopropyl β-D-thiogalactopyranose to a final concentration of 0.4 mM, cultures were incubated for an additional 12 h at 18 °C, and cells were harvested by centrifugation.

Cell pellets were resuspended in 30 ml of lysis buffer (25 mM sodium phosphate, pH 6.5, 300 mM NaCl, 20% glycerol, 5% glycerol), lysed by three passes through a French press at a pressure of 1000 pascals, and centrifuged at 15,000 rpm for 30 min at 4 °C. The supernatant was loaded onto a 5-ml HiTrap chelating column (GE Healthcare) pre-equilibrated with lysis buffer. The column was washed with 30 ml of Buffer A (25 mM Tris-HCl, pH 8.0, 300 mM NaCl, 5% glycerol) containing 50 mM imidazole. The protein was eluted with a linear gradient of 50–300 mM imidazole in Buffer A, first with and then without 2 mM EDTA, concentrated to 10 mg/ml using Centriprep-50 (Millipore), and stored at 20 °C. Yields for all recombinant forms of NS2B-NS3 protease were engineering proteins in which the 40-residue cofactor from NS2B (NS2B40, residues 54–93 of WNV NS2B) was fused to the N terminus of full-length NS3 (NS3FL) by a flexible 9-residue linker followed by a noncleavable linker Gly4-Ser-Gly4-His-Met. In a final construct, an inactive protease, Ser135-Ala, was created in NS2B40-G4SG4-HM-NS3FL/S135A. Gray boxes labeled H6 represent hexahistidine affinity tags. NS3 residues of the catalytic triad, His51 (H51), Asp75 (D75), and Ser135 (S135) are labeled, as is Gly1 (G1), the P1′ residue in the N-terminal cleavage site at the NS2B/NS3 junction.

Protease Assay—For self-cleavage reactions, purified proteins were diluted to the stated concentration (0.25–5.0 mg/ml) in assay buffer (25 mM Hepes, pH 8.5, 50 mM NaCl, 35% glycerol (v/v)), incubated at 37 °C for the indicated times, and quenched by addition of an SDS-PAGE loading buffer to a final concentration of 2% SDS. For intermolecular cleavage reactions, substrate and enzyme were diluted separately to 1 mg/ml in assay buffer, mixed in a 1:4 ratio (enzymesubstrate), unless indicated otherwise, and incubated for the indicated time.

 RESULTS

Design and Production of Recombinant Flavivirus NS2B-NS3—Flavivirus protease activity is dependent on the association of NS3 with a cofactor (NS2B40), a central 40-amino acid hydrophilic domain within the largely hydrophobic NS2B protein. To produce active NS2B-NS3 proteases from five flaviviruses, we engineered proteins in which the 40-residue cofactor from NS2B (NS2B40 and residues 54–93 of WNV NS2B) was fused to the N terminus of full-length NS3 (NS3FL) by a flexible 9-residue linker followed by a noncleavable dipeptide, Gly4-Ser-Gly4-His-Met (G4SG4-HM) (Fig. 2). A hexahistidine tag was engineered at the N terminus to facilitate purification of NS2B-NS3 protease. Similar recombinant forms of NS2B-NS3 protease were engineered by fusion of the respective NS2B cofactors with NS3 from YFV, DV2, DV4, and JEV. All recombinant proteases were produced in an E. coli expression system at 18 °C. The major proportion of the expressed protein was present in the soluble fractions of cell lysate, indicating that recombinant proteins were likely to be folded correctly. The proteins were purified by immobilized metal affinity chromatography via the His6 tags (in some cases followed by anion-exchange chromatography). SDS-PAGE analysis of purified proteins revealed fragments of lower molecular weight in addition to the expected products (Fig. 3, lanes 2, 5, 8, 11, and 14). These lower molecular weight products increased upon longer incubation at 37 °C (lanes 3, 6, 9, 12, and 15), indicative of proteolysis. However, none of the lower molecular weight fragments was detected in preparations of recombinant proteins in which the protease catalytic serine was substituted with alanine (Fig. 3, lanes 4, 7, 10, 13, and 16) and in which each protein was produced in identical yield to its corresponding parent. Thus, the recombinant NS2B-NS3 protease is subject to autolytic cleavage. At least two sites of autolysis are apparent in most of the NS2B-NS3 proteins.
Cis and Trans Activity of Flavivirus NS2B-NS3 Protease

Although the five proteases differ in their susceptibility to autolytic cleavage. Further analysis of cleavage was performed using the WNV NS2B$_{40}$G$_4$SG$_4$-HM-NS$_{3}$F$_{1}$ protein.

Sites of Autolytic Cleavage of WNV NS2B-NS3—Fragementutes of 8, 18.5, 50.5, and 69 kDa were detected in the recombinant NS2B-NS3 protein from WNV, indicative of two cleavage sites in the full-length protein of 77 kDa (Figs. 3 and 4). These sites were cleaved at different rates. One site was cleaved completely prior to purification, yielding fragments of 69 and 8 kDa (Fig. 4A, lane 2). Cleavage of the 69-kDa fragment at a second site was slower, as seen at time points during incubation at 37°C (Fig. 4A, lanes 2–6). All fragments were associated with the natively folded protein because they did not dissociate upon additional steps of nickel-affinity, anion-exchange, and size-exclusion chromatography of the monomeric protein (supplemental Fig. 1).

To identify the site of rapid cleavage, NS2B-NS3 was analyzed by mass spectrometry. A fragment of 7870 ± 8 Da was identified, which corresponds to cleavage at the “noncleavable” junction following the Gly$_4$-Ser-Gly$_4$-His-Met linker between NS2B and NS3 (compared with 7863 Da, calculated mass; data not shown). The cleavage site was also identified by N-terminal sequencing of the 69-kDa fragment, which yielded the sequence Gly-Gly-Val-Leu-Trp-Asp, corresponding to the N terminus of NS3, indicating that cleavage occurs at the site G$_4$S$_4$G$_4$HM ↓ GG.G.WD. Mutagenesis at this site, for experiments described below, provided further confirmation of the site identity. We designate this cleavage as an “N-terminal cleavage.” Cleavage at this site was unexpected because the sequence at the site, HM ↓ GG, is unlike the natural cleavage sequences at the critical P1 residue (Met versus Arg, see Fig. 1A), and substitution of Leu for Arg in the P1 position of the YFV NS2B-NS3 protease domain led to no detectable cleavage (29). In addition, autolytic cleavage of a similar NS2B$_{40}$-fused construct of WNV NS3$_{pro}$ was reported at NS3 residue Lys$_{15}$ (22), in a sequence (KK ↓ G) similar to natural flavivirus protease sites. We anticipated cleavage at Lys$_{15}$–Gly$_{16}$ when we detected the 8-kDa fragment, and we engineered a protein in which the nonconserved Gly$_{16}$ was substituted with Leu. This substitution had
no impact on appearance of the 8-kDa fragment (data not shown). We conclude that the Lys\textsuperscript{15}–Gly\textsuperscript{16} site is less accessible to the protease active site in full-length NS3 than it is in the isolated protease domain (NS\textsubscript{3\_pr}). We also attempted to isolate the full-length NS2B-NS3 protein by including protease inhibitors in the lysis buffer for cells from a fresh culture, but the protein was fully cleaved when it emerged from E. coli (supplemental Fig. 2).

The second, slower autolytic cleavage was presumed to occur within the helicase domain, since NS3 cleavage has been reported in cells infected with DV2 (10, 12), at a site corresponding to Arg\textsuperscript{459}–Gly\textsuperscript{460} in WNV NS3. Cleavage at this site was confirmed by N-terminal sequencing of the C-terminal fragment. The resulting sequence, Gly-Arg-Ile-Gly-Arg, corresponded to NS3 residues 460–464. The cleavage site was also confirmed by mutagenesis to produce the G460L substitution. The 8- and 69-kDa fragments were present in the purified proteins, but the 18.5- and 50.5-kDa fragments were not detected for NS2B\textsubscript{40}–G4SG4–HM–NS3\textsubscript{4/L} (Fig. 4C, lanes 14–18), confirming that the slower cleavage occurs in the helicase region at Arg\textsuperscript{459} → Gly\textsuperscript{460}. We designate this as the internal helicase case. No autolytic cleavages were observed with enzymatically inactive NS2B\textsubscript{40}–G4SG4–HM–NS3\textsubscript{4/L}/SI35A (Fig. 4B, lane 8–12).

An Artificial Substrate to Study Cleavage Specificity—To examine the unexpected N-terminal cleavage, we created an artificial substrate containing several variants of the natural NS2B–NS3 cleavage sequence (Fig. 1B). In this construct, a dodecapeptide encompassing the natural NS2B/NS3 junction was engineered between two fluorescent proteins, CFP and YFP, to create “CFP-2B/3-YFP” (LQYTKR\textsuperscript{P1} ↓ G\textsuperscript{P1} ↓ GVLWD). Proteolytic activity with this substrate was assayed by SDS-PAGE analysis of the reaction mixture, in which substrate (59.3 kDa, Fig. 5, lane 3), enzyme (lane 2), and two products (29.5 and 29.8 kDa, Fig. 5, lane 4) were easily distinguished, or by cleavage-induced loss of fluorescence resonance energy transfer between CFP and YFP (data not shown). We also made several artificial substrates containing substitutions in the NS2B/NS3 junction to check cleavage specificity. The artificial substrate (CFP-LQYTKR \downarrow GGVLWD–YFP) containing the natural sequence of the NS2B/NS3 junction was cleaved efficiently (Fig. 5, lane 4). Substitution of Asp for Arg at the P1 position (Fig. 5, lane 5) or Asp for Gly at the P1’ position (lane 6) abolished the cleavage activity. Two artificial substrates containing HM ↓ GG, the site that was cleaved rapidly in the full-length enzyme, were also tested. Neither CFP-LQYTHMGGLVLD–YFP, containing the natural NS2B/NS3 junction with KR replaced by HM, nor CFP-GGGHGMMGLVLD–YFP, containing the junction sequence from our recombinant enzyme, was cleaved in the artificial substrate (Fig. 5, lanes 7 and 8). This suggests that either HM ↓ GG is not a substrate for intermolecular cleavage or that a cleavage-competent conformation of HM ↓ GG is not accessible in the context of the artificial...
Cis and Trans Activity of Flavivirus NS2B-NS3 Protease

Variants of the parent enzyme (NS2B<sub>40</sub>-G<sub>4</sub>S<sub>4</sub>-H<sub>M</sub>-G<sub>1</sub>-NS3<sub>F1</sub>) were tested for autolytic cleavage activity (lanes 2–17). For each of the indicated enzyme variants, each protein is shown first as purified and second after a 2-h incubation at 37 °C. Except for the parent enzyme, all variants with non-natural sequences at the NS2B/NS3 junction were susceptible to autolysis at the helicase site but not at the N-terminal site (lanes 6–13). Variants with natural sequences at the NS2B/NS3 junction cleaved both sites (lanes 14–17). An artificial substrate (CFP-LQYTKR ↓GGVLWD-YFP, lane 18) was cleaved in trans by all variants (lanes 20–26) except the negative control (lane 19). The variant with the natural sequence at the NS2B/NS3 junction, NS2B<sub>40</sub>-G<sub>4</sub>S<sub>4</sub>-KR-G-NS3<sub>F1</sub>, had weak autolytic (presumably cis) activity at the helicase site (lane 15) and also weak trans activity (lane 25).

Substrate. The natural cleavage sites (Fig. 1B) for flavivirus NS2B-NS3 proteases generally have two basic residues (Arg-Arg, Arg-Lys or Lys-Arg) at the P2 and P1 positions and a small amino acid (Gly, Ser, or Ala) at the P1′ position. In contrast to the N-terminal HM ↓GG cleavage site, the internal helicase cleavage site has a sequence very similar to the natural inter-protein sites (Fig. 1B). We also made an artificial substrate, CFP-SAAQRR ↓GRIGRN-YFP, containing the Arg<sup>149</sup> ↓Gly<sup>360</sup> cleavage site observed in the recombinant protein. However, cleavage of this artificial substrate was barely detectable (Fig. 5, lane 9). In the context of the artificial substrate, efficient cleavage of the natural NS2B-NS3 site is in striking contrast to the lack of cleavage of the two sites of autolysis, suggesting that the autolytic reactions may occur in cis. Thus we designed materials to test this possibility explicitly.

Intramolecular Proteolytic Cleavage of WNV NS2B-NS3—To determine whether the autolytic cleavage is intramolecular (cis) or intermolecular (trans), we designed, expressed, and purified a natural substrate for trans cleavage (Fig. 6A). The protein, NS2B<sub>40</sub>-G<sub>4</sub>S<sub>4</sub>-HM-NS3<sub>F1</sub>/S135A, has intact N-terminal and helicase cleavage sites but lacks the protease catalytic side chain (Ser<sup>135</sup>). NS2B<sub>40</sub>-G<sub>4</sub>S<sub>4</sub>-HM-NS3<sub>F1</sub>/S135A undergoes no autolytic cleavage (Fig. 6, lane 3) in 2 h at 37 °C, whereas autolysis by the parent enzyme (NS2B<sub>40</sub>-G<sub>4</sub>S<sub>4</sub>-HM-NS3<sub>F1</sub>) results in complete cleavage at the N-terminal site and more than 50% cleavage at the internal helicase site under the same conditions (lane 9). We also designed, expressed, and purified an enzyme lacking both of the N-terminal and internal cleavage sites (NS2B<sub>40</sub>-G<sub>4</sub>S<sub>4</sub>-HD-NS3<sub>F1</sub>/G460L). This protein undergoes no autolytic cleavage (Fig. 6, lanes 6 and 7), as expected, but is an efficient enzyme against the artificial substrate CFP-LQYTKR ↓GGVLWD-YFP (lanes 17 and 18). To detect intermolecular hydrolysis, the substrate, NS2B<sub>40</sub>-G<sub>4</sub>S<sub>4</sub>-HM-NS3<sub>F1</sub>/S135A, was incubated with the enzyme, NS2B<sub>40</sub>-G<sub>4</sub>S<sub>4</sub>-HD-NS3<sub>F1</sub>/G460L. We observed no trans cleavage of substrate NS2B<sub>40</sub>-G<sub>4</sub>S<sub>4</sub>-HM-NS3<sub>F1</sub>/S135A (Fig. 6, lanes 10 and 11) by enzyme NS2B<sub>40</sub>-G<sub>4</sub>S<sub>4</sub>-HD-NS3<sub>F1</sub>/G460L under conditions in which this enzyme completely hydrolyzed the artificial substrate (lanes 17 and 18). These results clearly indicate intramolecular cleavage of both the N-terminal site (GGGGHM ↓GGVLWD) and the internal helicase site (SAAQRR ↓GRIGRN) of recombinant NS2B-NS3. The natural cleavage sequence at the NS2B/NS3 junction was also tested for trans cleavage by construction of the substrate NS2B<sub>40</sub>-G<sub>4</sub>S<sub>4</sub>-KR-NS3<sub>F1</sub>/S135A. A low level of N-terminal cleavage was seen using the trans substrate (Fig. 6, lanes 12 and 13), but this was far below the rate of cis cleavage of the parent construct (lanes 8 and 9). Further confirmation of cis cleavage was provided by a dilution experiment using purified NS2B<sub>40</sub>-G<sub>4</sub>S<sub>4</sub>-HM-NS3<sub>F1</sub>. The rate of cleavage of the Arg<sup>149</sup> ↓Gly<sup>360</sup> helicase site was independent of a 20-fold difference in protease concentration, as expected for an intramolecular reaction (Fig. 6B).

Comparison of Cis and Trans Cleavage Activities in Mutant Proteases—We next examined how a variety of sequences at the NS2B-NS3 cleavage site in a natural substrate affect autolytic cleavage and trans cleavage of an artificial substrate (Fig. 7). Several enzymes were made with mutations at the P1 and P1′ positions of the N-terminal cleavage site (G<sub>4</sub>S<sub>4</sub>GGGHHM<sup>110</sup> ↓G<sub>1</sub>-G<sub>1</sub>-GVVLWD). Substitution of a negatively charged side chain (Asp) at the P1 site or P1′ site, or substitution of a large side chain (Leu or Trp) at the P1′ site, eliminated autolytic cleavage at the NS2B/NS3 junction (Fig. 7, lanes 6–13). All enzymes that were not susceptible to N-terminal autolytic cleavage at the NS2B/NS3 junction had efficient trans cleavage activity with the artificial trans substrate, CFP-LQYTKR ↓GGVLWD-YFP (lanes 21–24), and efficient cis cleavage of the internal helicase site (Fig. 7, lanes 7, 9, 11, and 13). This demonstrates that prote-
ase activity requires neither cleavage of NS2B from NS3 nor a free NS3 N terminus. Previous results have been mixed in this regard. Pugachev et al. (30) suggested that NS2B–NS3 cleavage is required to release NS3 for internal cleavage at the helicase site. However, Teo and Wright (12) demonstrated that prior cleavage between NS2B and NS3 was not necessary for NS2B–NS3 protease activity.

NS2B<sub>40</sub>-G<sub>4</sub>S<sub>4</sub>-KR-NS3<sub>FL</sub>, with the wild type sequence at the NS2B/NS3 junction, had the expected autolytic cleavage at the N-terminal site (Fig. 7, lane 14). However, compared with the NS2B<sub>40</sub>-G<sub>2</sub>S<sub>2</sub>-HM-NS3<sub>FL</sub>, parent, NS2B<sub>40</sub>-G<sub>2</sub>S<sub>2</sub>-KR-NS3<sub>FL</sub>, exhibited unexpectedly low levels of cleavage at the helicase site (Fig. 7, compare lanes 5 and 15) and with the trans substrate CFP-LQYTKR ↓ GGVLWD-YFP (compare lanes 20 and 25). This is in contrast to substitution of Gln for Arg at the cleavage site, which resulted in efficient cis (Fig. 7, lane 17) and trans cleavage (lane 26). The simplest explanation of these data is a form of product inhibition, in which the reaction product Lys-Arg-COO<sup>−</sup> remains in the P2-P1 site after cleavage and effectively blocks the active site for both cis and trans substrates.

Effect of Linker Length on Catalytic Activity of NS2B–NS3—We next examined whether Gly<sub>4</sub>Ser-Gly<sub>4</sub> tethering the NS2B cofactor to NS3, was responsible for the unexpectedly low cleavage activity of NS2B<sub>40</sub>-G<sub>2</sub>S<sub>2</sub>-KR-NS3<sub>FL</sub>. The effect of the length of the NS2B<sub>40</sub>-NS3 linker on protease activity was studied in proteins having linkers longer and shorter than the Gly<sub>4</sub>Ser-Gly<sub>4</sub> linker of the parent protein, NS2B<sub>40</sub>-G<sub>2</sub>S<sub>2</sub>-HM-NS3<sub>FL</sub> (Fig. 8). We engineered NS2B<sub>79</sub>-NS3<sub>FL</sub>, in which the 38 residues of the natural C terminus of NS2B replaced Gly<sub>4</sub>Ser-Gly<sub>4</sub>-HM and were fused to NS3 as in NS3<sub>FL</sub>, in which the Gly<sub>4</sub>Ser-Gly<sub>4</sub> linker was absent and the natural substrates of WNV NS2B–NS3 protease. The fusion substrate (CFP-LQYTKR ↓ GGVLWD-YFP) is a form of product inhibition, in which the reaction product Lys-Arg-COO<sup>−</sup> remains in the P2-P1 site after cleavage and effectively blocks the active site for both cis and trans substrates.

DISCUSSION

This is the first study to examine the activity of the WNV NS2B<sub>40</sub>-NS3 protease in the context of purified, full-length NS3 protein. Previous detailed biochemical studies of flavivirus proteases were of truncations (13–16, 19–22) that included only the protease region of NS3 (WNV NS3 residues 1–180) but lacked the full helicase region (WNV NS3 residues 180–619). The parent protein for our studies, NS2B<sub>40</sub>-G<sub>4</sub>S<sub>4</sub>-HM-NS3<sub>FL</sub>, was a fusion of the 40-residue cofactor of NS2B and full-length NS3. A nonapeptide linker (Gly<sub>4</sub>Ser-Gly<sub>4</sub>) was selected based on previously reported active fusions of NS2B and NS3 protease domains from DV2 and WNV (17–22). This linker was connected to the N terminus of NS3 with the dipetide His-Met, which differs from sequences (Lys-Arg) of the natural substrates of WNV NS2B–NS3 protease. The fusion product, NS2B<sub>40</sub>-G<sub>4</sub>S<sub>4</sub>-HM-NS3<sub>FL</sub>, was a stable, monomeric protein, in contrast to full-length NS3 in absence of the NS2B cofactor, which exhibited nonspecific aggregation and poor solubility. This is consistent with the crystal structure of the NS2B–NS3 protease domain, in which the NS2B cofactor buries several hydrophobic surface patches as it wraps around the NS3 protease domain (Fig. 9) (24). The importance of hydrophobic contacts between NS2B and the NS3 protease domain was illustrated in a recent site-directed mutagenesis study of the proteins (21).

NS2B<sub>40</sub>-G<sub>4</sub>S<sub>4</sub>-HM-NS3<sub>FL</sub>, the parent enzyme for our studies, was susceptible to autolytic cleavage at two sites (Fig. 4). All cleavage fragments remained associated with the protein, as seen by the identical behavior of cleaved and uncleaved proteins upon ion-exchange and gel filtration chromatography (supplemental Fig. 1). The autolytic cleavage reactions were intramolecular (Fig. 6). Although these autolytic reactions were characterized in detail only for WNV NS2B–NS3 protease, they also occurred to varying extents in the YFV, DV2, DV4, and JEV proteases (Fig. 3), where we assume they are also intramolecu-
The internal helicase cleavage site (SAAQRR↓GRIGRN) is similar to the other WNV cleavage sites (Fig. 1B), and we were surprised that it was not cleaved in trans. The main difference to other natural cleavage sites is Arg⁴⁶¹ at the P₂' position. Lack of trans cleavage of this site implies some degree of protease specificity at the P₂' position. Other WNV protease sites have Gly, Gln, or Trp at the P₂' position, although it is unknown whether most of them are cleaved in cis or trans. Indirect evidence suggests that the NS2A-NS2B site (P₂' Trp) is cleaved in cis (5), and the structure of the helicase domain suggests that the NS3-NS4A site is inaccessible for cis cleavage and is therefore cleaved in trans (25). On this basis, charge rather than size may be the problem with trans cleavage of substrates having P₂' Arg.

The cleavage of the noncognate His-Met ↓ Gly at the N-terminal site is attributed to the high effective concentration of a cis substrate and to the Gly₄-Ser-Gly₄ linker, which must facilitate binding of His-Met in the protease P₂-P₁ sites. The situation in the viral polyproteins is more complex. In extensive studies of YFV protease specificity, Rice and co-workers (11, 29, 31) reported high specificity (Arg and Lys) at the P₁ position for the NS2A/NS2B and NS2B/NS3 junctions, but more relaxed specificity at the NS3:NS4A (Arg, Lys, Ser, Thr, Ala, Met, and Leu) and NS4B:NS5 (Arg, Lys, Gln, Asn, and His) junctions.

We detected no cleavage at Lys¹⁵–Gly¹⁶ of NS3, as was observed in the variant of the isolated WNV NS2B-NS3 protease domain used for the crystal structure (24). This is easily explained by the structure, in which the first (Thr¹⁹) and last (Arg⁷⁷⁰) ordered residues of the NS₃_pro domain are less than 20 Å apart. Presumably, the presence of the larger helicase domain (residues 180–619) in full-length NS3 protected Lys¹⁵–Gly¹⁶ from the protease active site.

Our results show that neither cleavage of the NS2B/NS3 junction nor a free NS3 N terminus is required for protease activity. Variants of the parent protein in which cis cleavage of NS2B from NS3 was blocked were effective both in trans cleavage reactions and in cis cleavage of the helicase site (Fig. 7).

Cis cleavage at the N-terminal site of our recombinant proteins depended upon the sequence at the cleavage site and also on the length of the linker between NS2B₄₀ and NS₃₇₁. The results are easily rationalized by the structure of the WNV NS2B-NS3 protease domain bound to a product analog (24), which shows how the dipeptide Arg-Arg binds in the protease P₂-P₁ sites (Fig. 9). The distance of peptide...
travel would be minimally 20 Å from the last residue of NS2B in the crystal structure (residue 88, labeled C in Fig. 9) to the residue in the protease P2 site. In our recombinant proteins, this corresponds to 14 residues in NS2B40-G4SG4-HM-NS3FL and NS2B40-G4SG4-KR-NS3FL. However, it corresponds to only 5 residues in NS2B40-HM-NS3FL and NS2B40-KR-NS3FL, too few residues to span 20 Å. In recombinant proteins with a 14-residue connector, both noncognate (His-Met) and cognate (Lys-Arg) dipeptides bound productively in the P2-P1 sites, and both NS2B40-G4SG4-HM-NS3FL and NS2B40-G4SG4-KR-NS3FL were cleaved in cis. However, with the shorter 5-residue connector, only the cognate dipeptide (Lys-Arg) bound productively, so that NS2B40-KR-NS3FL was not associated with a biological membrane. Based on near the C terminus of NS2B40, which form part of the P2 substrate site (Fig. 9). It may be impossible to form this part of the P2 site and simultaneously place the 10th downstream residue into the P2 site, as would be required with a 5-residue connector. In contrast, the P1 site is formed by NS3 alone (Fig. 9), and binding of cis substrates in the P1 site may lead to cleavage even without binding in the P2 site. In NS2B40-KR-NS3FL, Arg could reach the P1 site, perhaps at the expense of disassembling the P2 site, leading to cleavage. However, in NS2B40-HM-NS3FL, Met had no intrinsic affinity for the P1 site and was not cleaved. The structure interpretation also explains why NS2B40-G4SG4-KR-NS3FL was a poor enzyme in our studies. The longer 14-residue connector allowed the cleavage product Lys-Arg-COO" to remain bound in the P2-P1 sites, effectively blocking other substrates from entering, whereas the shorter 5-residue connector did not allow the product to remain in the active site. Thus, NS2B40-KR-NS3FL was a better enzyme than was NS2B40-G4SG4-KR-NS3FL.

The apparent product inhibition by the cognate Lys-Arg dipeptide at the NS2B/NS3 junction of NS2B40-G4SG4-KR-NS3FL raises the possibility that slow dissociation of the NS2B C terminus from the protease active site may provide a useful pause in viral polyprotein processing. For example, slow dissociation could allow time for downstream events, such as protein folding or protein-protein associations, to occur before further polyprotein processing. Our recombinant NS2B-NS3 proteases differ from the natural situation in two critical ways. First, the NS2B/NS3 junction in our studies was very close to the end of the NS2B cofactor and not separated by the remaining 35 residues of NS2B. No product inhibition was observed in the construct with the full C terminus of NS2B (Fig. 8). Second, in our experiments NS2B was not associated with a biological membrane. Based on predicted membrane domains within NS2B, the distances from the NS2B cofactor to the ER membrane are rather short, ~3 residues from the cofactor N terminus and ~12 residues from the C terminus. If the membrane-associated predictions are correct, then the NS3 protease domain hangs from the ER membrane in a “sling” formed by the NS2B cofactor (Fig. 9). This geometry places the protease active site rather near the ER membrane, no matter how flexible the tethers. This is consistent with positions of the viral protease cleavage sites, which are all proximal to predicted membrane domains of prM, NS2A, NS2B, NS4A, and NS4B.

The results presented here provide new insights into the function of the flavivirus NS3 protease. Differences in the catalytic profile of the protease in full-length NS3 compared with the isolated protease domain are important for protease function in a virus-infected cell.

Acknowledgments—We thank Richard Kinney for kindly providing plasmids encoding the WNV NS2B-NS3, DV2 NS2B-NS3, and DV4 NS2B-NS3; Charles Rice for plasmid encoding the YFV NS2B-NS3; Tsutomu Takegami for providing a partial cDNA clone of JEV; and Todd Geders for providing plasmid pCYFP28.

REFERENCES


Cis and Trans Activity of Flavivirus NS2B-NS3 Protease
Cis and Trans Activity of Flavivirus NS2B-NS3 Protease