Site-directed Mutagenesis and Kinetic Studies of the West Nile Virus NS3 Protease Identify Key Enzyme-Substrate Interactions*

Received for publication, August 30, 2004, and in revised form, September 27, 2004
Published, JBC Papers in Press, October 19, 2004, DOI 10.1074/jbc.M409931200

Keith J. Chappell‡, Tessa A. Nall§, Martin J. Stoormer‡, Ning-Xia Fang‡, Joel D. A. Tyndall§§, David P. Fairlie§, and Paul R. Young¶¶

From the ‡Department of Microbiology and Parasitology, School of Molecular and Microbial Sciences, University of Queensland, Brisbane, Queensland 4072, Australia, §Centre for Drug Design and Development, Institute for Molecular Bioscience, University of Queensland, Brisbane, Queensland 4072, Australia, and ¶¶School of Pharmacy, University of Otago, P. O. Box 913, Dunedin 9015, New Zealand

The flavivirus West Nile virus (WNV) has spread rapidly throughout the world in recent years causing fever, meningitis, encephalitis, and fatalities. Because the viral protease NS2B/NS3 is essential for replication, it is attracting attention as a potential therapeutic target, although there are currently no antiviral inhibitors for any flavivirus. This paper focuses on elucidating interactions between a hexapeptide substrate (Ac-KPGLKR-p-nitroanilide) and residues at S1 and S2 in the active site of WNV protease by comparing the catalytic activities of selected mutant recombinant proteases in vitro. Homology modeling enabled the predictions of key mutations in WNV NS3 protease at S1 (V115A/F, D129A/E/N, S135A, Y150A/F, S160A, and S163A) and S2 (N152A) that might influence substrate recognition and catalytic efficiency. Key conclusions are that the substrate P1 Arg strongly interacts with S1 residues Asp-129, Tyr-150, and Ser-163 and, to a lesser extent, Ser-160, and P2 Lys makes an essential interaction with Asn-152 at S2. The inferred substrate-enzyme interactions provide a basis for rational protease inhibitor design and optimization. High sequence conservation within flavivirus proteases means that this study may also be relevant to design of protease inhibitors for other flavivirus proteases.

West Nile virus (WNV) is a member of the Flavivirus genus and is transmitted by mosquitoes, primarily Culex species (1), between avian reservoir hosts and vertebrate dead-end hosts including humans and horses. Many viruses within this genus are medically important pathogens, including dengue, Japanese encephalitis, tick-borne encephalitis, and yellow fever viruses. WNV was first isolated in 1937 in Uganda’s West Nile province and was subsequently found in regions of Africa, the Middle East, Europe, Russia, Southwestern Asia, and Australia (less severe subtype, Kunjin) (2). Human infection is generally asymptomatic or causes a mild febrile disease, West Nile fever (1). However, in a small number of cases, predominantly in the elderly, the infection with WNV results in encephalitis or meningitis that can be fatal (1). Over the last decade, there has been an increase in the frequency of human outbreaks and severity of disease with recent epidemics in Israel (1998), Romania (1999), Russia (1999), and New York (1999) (3). Since the introduction of WNV into New York in 1999, it has spread rapidly throughout the United States (4,156 infections and 284 deaths in 44 states in 2002, 9862 infections and 264 deaths in 2003) (reported in 2000 by the CDC, Division of Vector-Borne Infectious Disease: West Nile Virus, Center for Disease Control and Prevention, www.cdc.gov/ncidod/dvbid/westnile/index.htm), Canada, and Mexico and has recently appeared in the United Kingdom (5). There is currently no vaccine or antiviral treatment available for human WNV infection. However, a chimeric live vaccine is in clinical trials and a veterinary vaccine is licensed for use in equines and exotic zoo birds (6).

Flaviviruses are small enveloped viruses containing a single-stranded positive sense RNA genome of 10–11 kb with a single large open reading frame encoding a polypeptide precursor of ~3400 amino acids. Gene expression requires both host and a virally encoded protease to process the polypeptide precursor into the individual functional proteins. They comprise three structural proteins (C, prM, and E) and seven non-structural proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5) (7). The viral protease encoded within the N-terminal third of NS3 is responsible for cleavage at the NS2A/NS2B, NS2B/NS3, NS3/NS4A, and NS4B/NS5 gene junctions and also at a site near the C terminus of the C protein to promote efficient generation of prM (Fig. 1) (8, 9). Mutation of residues at any of these cleavage sites in the related yellow fever virus was shown to prevent efficient cleavage and abolished virus infectivity in cell culture (10–13), highlighting a vital role for the NS3 protease in replication. The essential nature of this protease in the virus life cycle is the basis for interest in NS3 as a possible target for developing antiviral inhibitors.

NS3 is a multifunctional protein in which the N-terminal 184 amino acids encode for the protease and the C-terminal region encodes a nucleotide triphosphatase, an RNA triphosphatase, and a helicase (14–16). The NS3 protease is a trypsin-like serine protease with a classic catalytic triad (His-51, Asp-75, and Ser-135) (17). Protease activity has been shown for a number of related flaviviruses to be dependent on the association of NS2B as a cofactor (18). A central 40 amino acid hydrophilic domain within the largely hydrophobic NS2B protein has been shown to be sufficient for cofactor activity (19, 20). The flanking hydrophobic domains within NS2B are likely to function in promoting membrane association of NS2B-NS3 (21, 22). NS2B-NS3pro has high specificity for substrate processing requiring a dibasic recognition sequence

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† To whom correspondence should be addressed. Fax: 61-733654620; E-mail: p.young@uq.edu.au.

‡ The abbreviations used are: WNV, West Nile virus; pNA, p-nitroanilide; HCV, hepatitis C virus; CHAPS, 3-(3-cholamidopropyl)dimethylammonio-1-propanesulfonic acid; R	extsubscript{t}, retention time.
(P2-Lys, P1-Arg) that is conserved throughout the flaviruses (7). This relatively unusual substrate specificity suggests that inhibitors could be somewhat selective over most host serine proteases (19).

Fundamental to the design of a specific inhibitor is a detailed understanding of the interactions between protease residues in the active site and substrates. The structure of the WNV NS3 protease is unknown, but there are reported crystal structures for related NS3 proteases of hepatitis C virus (HCV) with a cofactor (23) and Dengue-2 virus without NS2B cofactor (24). Therefore, the latter structure was of an inactive enzyme. In this study, a homology-modeled structure of the WNV NS3 protease, created through sequence comparison with the HCV and Dengue-2 virus NS3 protease structures, was used to predict residues in the S1 and S2 binding pockets that probably make important interactions with substrate residues (25). Identified residues were then mutated in a recombinant catalytically active construct by site-directed mutagenesis, and kinetic parameters for the mutant WNV proteases were determined and compared.

**Experimental Procedures**

**General Methods**—Protected amino acids and resins were obtained from Auspep, Novabiochem, and PepChem. Trifluoroacetic acid, piperidine, N,N-diisopropylamine, dichloromethane, and N,N-dimethylformamide (peptide synthesis grade) were purchased from Auspep. All of the other materials were of reagent grade unless otherwise stated. Crude peptides were purified by reversed-phase high pressure liquid chromatography on a Vydac C18 column (10–15 μm, 300 Å, 50 × 250 mm) using a gradient mixture of solvent (A) 0.1% trifluoroacetic acid/water and (B) 0.1% trifluoroacetic acid/10% water/90% acetonitrile.

Analytical reversed-phase high pressure liquid chromatography was performed on a Waters system equipped with a 717 plus autosampler, 660 controller, and a 996 photodiode array detector using a reversed-phase Phenomenex Luna C18 column (5 μm, 100 Å, 250 × 4.6 mm). Purified peptides were characterized by analytical reversed-phase high pressure liquid chromatography (linear gradient 0–100% B over 30 min), mass spectrometry, and 1H NMR spectroscopy. The molecular mass of the peptides was determined by electrospray mass spectrometry on a Micromass LCT mass spectrometer. 1H NMR spectra were recorded on samples containing 4 mM peptide in Me2SO-d6 (550 μl) on a Bruker Avance 600 spectrometer at 298 K. Proton assignments were determined by TOCSY (80 ms of mixing time), DQF-COSY, ECOSY, and NOESY (350 ms of mixing time) spectra using the sequential assignment method (26). All of the spectra were processed on Silicon Graphics R10000 or R12000 workstations using XWINNMR, version 2.6 (27).

**Homology Modeling**—West Nile virus NS3 protease homology models were generated using the structures of Dengue-2 virus NS3 protease without cofactor (Protein Data Bank 1BEF) (24) and with Bowman-Birk inhibitor (Protein Data Bank 1DF9, subunit B) (29) and hepatitis C virus NS3 protease with bound cofactor (Protein Data Bank 1AIQ) (23). Models were generated using the Modeler and Homology modules within InsightII (31) on a Silicon Graphics R10000 workstation. Sequences were aligned using Align 2D, Structure Alignment (Homology module, InsightII), and ClustalW alignment (32). Secondary structure predictions were conducted using the web-based PsiPred prediction for each of the three NS3 protease sequences (33). Electrostatic potential mapping was performed on the WNV NS3 protease model using the Delphi module in InsightII (34). A four residue P3-P1′ substrate based on the NS4B-NS5 cleavage site (Leu-Lys-Arg-Gly) was docked into the active site of a homology model of the WNV NS3 protease using GOLD, version 2.1 (35). A specified distance constraint between the substrate residue Arg (P1) and WNV protease Tyr-150 (S1) was used based on our

![Diagram](http://www.jbc.org/Downloaded from)

**FIG. 1.** Flavivirus polyprotein processing and site-directed mutagenesis constructs. A, the upper schematic shows sites cleaved by host proteases (shaded arrows) and the virus-encoded NS2B/NS3 protease (open arrows). The proteolytic domain of NS3 and the NS2B cofactor are shaded. The schematic below shows this region expanded and the recombinant construct used in this study. Features of the NS2B/NS3pro complex include the cleavage site (open arrow), the catalytic triad (asterisks), the hydrophobic domains of NS2B (shaded arrows), the cofactor domain required for catalytic activity. The construct CF40.Gly.NS3pro comprises this essential 40 residue domain of NS2B linked to the proteolytic domain of NS3 and the NS2B cofactor are depicted residues important for related NS3 proteases of hepatitis C virus (HCV) with a cofactor (23) and Dengue-2 virus without NS2B cofactor (24). Therefore, the latter structure was of an inactive enzyme. In this study, a homology-modeled structure of the WNV NS3 protease, created through sequence comparison with the HCV and Dengue-2 virus NS3 protease structures, was used to predict residues in the S1 and S2 binding pockets that probably make important interactions with substrate residues (25). Identified residues were then mutated in a recombinant catalytically active construct by site-directed mutagenesis, and kinetic parameters for the mutant WNV proteases were determined and compared.

![Diagram](http://www.jbc.org/Downloaded from)
understanding of protease-inhibitor interactions (36, 37). 2 Docked structures were ranked for best-fitting conformations using GOLD, version 2.1 (35).

Enzyme Expression and Purification—The pQE9 vector was used for high level inducible expression of N-terminal hexahistidine-tagged recombinant proteins. Cultures of the Escherichia coli strain G13009 transformed with the expression plasmid were grown in 500 mL of LB medium containing 100 μg/ml ampicillin and 25 μg/ml kanamycin at 37 °C until the A600 reached 0.5. The expression of the recombinant protein was induced by the addition of isopropyl-β-D-thiogalactopyranoside to a final concentration of 0.3 mM and incubated for an additional 3 h at 22 °C. Cells were harvested by centrifugation and stored at −20 °C.

For protein purification, cell pellets were thawed, resuspended in lysis buffer (5 mL/g wet pellet, 50 μM HEPES, pH 7.5, 300 mM NaCl, 10 mM imidazole, 5% glycerol) and protease inhibitors were added (1 μg/mL aprotinin, 1 μg/mL leupeptin, 1 μg/mL benzamidine, 1 mM phenylmethylsulfonyl fluoride, 10 μg/mL TACL, all final concentrations) in an attempt to minimize proteolytic cleavage of recombinant protein. Subsequent analyses have shown that these standard serine protease inhibitors do not inhibit WNV CF40.Gly.NS3pro (25). Resuspended cells were lysed by three passages through a French press (1000 Pascals) and then centrifuged (27,000 × g, 20 min, 4 °C).

The recombinant proteases, each with a N-terminal hexahistidine tag, were purified by affinity chromatography on a Ni2+-nitrilotriacetic acid-agarose 1-cm column (Qiagen) that had been pre-equilibrated (50 mL HEPES buffer (30 mM), pH 7.5, 300 mM NaCl, 10 mM imidazole, 5% glycerol). Resin was then removed, mixed with the supernatant fraction of cell lysates, and incubated at 4 °C on a rotator for 30 min to allow the His-tagged protein to bind to the Ni2+ resin. The column was washed, and washed with 30 mL of buffer containing 20 mM imidazole, and the proteins were eluted into 6 × 1-mL fractions with buffer containing 100 mM imidazole. The pre- and post-isopropyl-β-D-thiogalactopyranoside induced samples, soluble and insoluble fractions, and the elution fractions were analyzed by 12% SDS-PAGE.

**Table I**

| Primer         | Sequence                                                                 |
|---------------|--------------------------------------------------------------------------|
| WNV CF40a.F   | 5'-CGATGACCGCGCCAGGATCACAGATATGGGA TTGAGAGAACG-3'                      |
| WNV NSpro.R   | 5'-GCCCGCAAGCTTACAGCATCTCGAGTGGAT-3'                                    |
| WNV NS3_V115A.F | 5'-ACGCAAACAGAGCCTGCAAACACCT-3'                                         |
| WNV NS3_V115A.R | 5'-AGGTGCTTTGAGGTGGTTTCTGCTG-3'                                        |
| WNV NS3_D129F | 5'-ACGAAACAGAGCCTGCAAACACCT-3'                                         |
| WNV NS3_D129R | 5'-AGGTGCTTTGAGGTGGTTTCTGCTG-3'                                        |
| WNV NS3_S135A.F | 5'-TTCCAGTGGAAGATCCCAAGTGACG-3'                                       |
| WNV NS3_S135A.R | 5'-GCCCTGACCTTGCGTCTGAACTGGC-3'                                      |
| WNV NS3_Y150A.F | 5'-GCCCGCAAGCTTACAGCATCTCGAGTGGAT-3'                                   |
| WNV NS3_Y150A.R | 5'-AGGTGCTTTGAGGTGGTTTCTGCTG-3'                                        |
| WNV NS3_Y150F | 5'-GCCCGCAAGCTTACAGCATCTCGAGTGGAT-3'                                   |
| WNV NS3_Y150R | 5'-AGGTGCTTTGAGGTGGTTTCTGCTG-3'                                        |
| WNV NS3_Y150F.R | 5'-TTCCAGTGGAAGATCCCAAGTGACG-3'                                      |
| WNV NS3_Y150R.R | 5'-GCCCTGACCTTGCGTCTGAACTGGC-3'                                      |
| WNV NS3_Y150F | 5'-GCCCGCAAGCTTACAGCATCTCGAGTGGAT-3'                                   |
| WNV NS3_Y150R | 5'-AGGTGCTTTGAGGTGGTTTCTGCTG-3'                                        |
| WNV NS3_Y150F.R | 5'-TTCCAGTGGAAGATCCCAAGTGACG-3'                                      |
| WNV NS3_Y150R.R | 5'-GCCCTGACCTTGCGTCTGAACTGGC-3'                                      |
| WNV NS3_Y150F | 5'-GCCCGCAAGCTTACAGCATCTCGAGTGGAT-3'                                   |
| WNV NS3_Y150R | 5'-AGGTGCTTTGAGGTGGTTTCTGCTG-3'                                        |
| WNV NS3_Y150F.R | 5'-TTCCAGTGGAAGATCCCAAGTGACG-3'                                      |
| WNV NS3_Y150R.R | 5'-GCCCTGACCTTGCGTCTGAACTGGC-3'                                      |

*Tyndall, J. D. A., Nall, T., and Fairlie, D. P. (2005) Chem. Rev., in press.*
eight different substrate concentrations ranging from 62.5 μM to 1 mM for determination of kinetic constants. To obtain accurate kinetic data for the D129A site-directed mutant, which possessed only slightly higher activity than the negative control, it was assayed in quadruplicate at the higher enzyme concentration of 2.5 μM and eight higher substrate concentrations ranging from 125 μM to 2 mM. Kinetic parameters were calculated from weighted non-linear regression of the initial velocities as a function of the eight substrate concentrations using GraphPad Prism 4 software. The $k_{cat}/K_m$ values were calculated assuming that Michaelis-Menten kinetics $v = V_{max}[S]/(K_m + [S])$. Triplicate measurements were taken for each data point. The data are reported as the means ± S.E.

RESULTS

**Protease Architecture and Design of Mutant Proteases**—A structural homology model of the WNV NS3pro (25) was reconstructed here based on the sequence alignments and the crystal structures of Dengue-2 virus NS3pro in complex with a Bowman-Birk inhibitor (29) and of HCV NS3pro in complex with its cofactor NS4A (Fig. 2) (28). The HCV NS3pro structure was included to enable the modeling of the WNV NS3pro in the context of its NS2B cofactor.

The electrostatic surface map of WNV NS3pro is shown in Fig. 2A, highlighting acidic (red), basic (blue), and hydrophobic (white) regions with a previously reported substrate (P3-P1' segment, Leu-Lys-Arg-Gly) shown in green (stick model). B, enlarged view of the substrate bound in the protease active site. Two acidic residues, the catalytic Asp-75 as well as Asp-129 at the rim of the deep P1 binding pocket (S1 subsite), are highlighted. The P3 to P1' residues are also identified. C, magnified view of substrate binding groove (substrate in gold) showing putative substrate-interacting residues in the S1 (Val-115, Asp-129, Tyr-150, Ser-160, and Ser-163) and S2 (Asn-152) subsites that were targeted for site-directed mutagenesis. D, substrate-WNV protease interactions (substrate in gold and protease in blue). Catalytic Ser-135 residue is displayed in bold and denoted with an asterisk. Protease substrate regions are shown in boldface (S1, S2, and S1') along with residues predicted by the model to interact with the substrate.

**FIG. 2.** WNV NS3pro homology model. A, electrostatic surface map highlighting acidic (red), basic (blue), and hydrophobic (white) regions with a bound substrate (P3-P1' segment, Leu-Lys-Arg-Gly) shown in green (stick model). B, enlarged view of the substrate bound in the protease active site. Two acidic residues, the catalytic Asp-75 as well as Asp-129 at the rim of the deep P1 binding pocket (S1 subsite), are highlighted. The P3 to P1' residues are also identified. C, magnified view of substrate binding groove (substrate in gold) showing putative substrate-interacting residues in the S1 (Val-115, Asp-129, Tyr-150, Ser-160, and Ser-163) and S2 (Asn-152) subsites that were targeted for site-directed mutagenesis. D, substrate-WNV protease interactions (substrate in gold and protease in blue). Catalytic Ser-135 residue is displayed in bold and denoted with an asterisk. Protease substrate regions are shown in boldface (S1, S2, and S1') along with residues predicted by the model to interact with the substrate.
and processing. The S160A and S163A mutants were made to test the possible hydrogen-bonding interaction of these residues in S1 with Arg at P1 of the substrate.

Asp-129 is highly conserved among the flaviviruses and corresponds to Asp-189 of trypsin, a critical residue that sits at the base of the substrate binding pocket and forms an electrostatic bond with the P1 Lys or Arg of the trypsin cleavage site (30). However, as noted above, Tyr-150 takes up this position in the flavivirus NS3 protease with Asp-129 located more peripherally at the outer edge of the S1 site (see Fig. 2A). Despite its high level of conservation across the flaviviruses, previous cell-based mutagenesis studies appeared to show that Asp-129 could be relatively freely substituted with retention of significant activity (38). In addition, crystallographic data for the Dengue-2 virus NS3 suggest that Asp-129 does not appear to interact with the P1 Lys but to only interact with a P1 Arg in one of two conformations (29). We decided to test more accurately the importance of Asp-129 for substrate processing by making three mutant WNV CF40.Gly.NS3pro enzyme constructs (D129A, D129E, and D129N). Conservative mutants D129E (no change in charge) and D128N (isosteric replacement but loss of charge) together with D129A (loss of charge and bulk) at S1 would clarify whether a charged group is important for interaction with Arg at P1 of the substrate. The mutations of V115F and V115A were made to clarify the importance of Tyr-150 and determine the actual size of the S1 pocket. The rationale behind these two mutations was that V115F could fill the space available for the binding of a basic residue while being complimentary to the existing Tyr-150. V115A was designed to ascertain the importance of valine versus the size of the cavity.

Because of the known dibasic P1/P2 substrate specificity of WNV NS3 protease, we also wanted to make the mutations at the S2 site in the enzyme. The inspection of the homology structure suggested that Asn-152 was most likely to hydrogen bond to the basic P2 Lys residue of the substrate, so we chose the N152A mutation to test this effect. An inactive protease was constructed by mutation of the catalytic serine 135 to alanine as a negative control.

Comparative kinetic studies of these mutant proteases and the assessment of the role of the mutated residues in substrate binding were expected to validate the predictions made from the homology structure regarding enzyme-substrate interactions.

Expression and Purification of WNV CF40.Gly.NS3pro and Mutants—The recombinant WNV protease (WNV CF40.Gly.NS3pro) used in this study comprises the essential 40 amino acid cofactor domain of NS2B joined to the 184 amino acid protease domain of NS3 by a flexible nonapeptide G4SG4,a b-amidase cofactor domain of NS2B joined to the 184 amino acid protease domain of NS3 by a flexible nonapeptide G4SG4,a b-amidase domain (25). Comparative kinetic studies of these mutant proteases and the assessment of the role of the mutated residues in substrate binding were expected to validate the predictions made from the homology structure regarding enzyme-substrate interactions.

Expression and Purification of WNV CF40.Gly.NS3pro and Mutants—The recombinant WNV protease (WNV CF40.Gly.NS3pro) used in this study comprises the essential 40 amino acid cofactor domain of NS2B joined to the 184 amino acid protease domain of NS3 by a flexible nonapeptide G4SG4,a b-amidase domain (25). Comparative kinetic studies of these mutant proteases and the assessment of the role of the mutated residues in substrate binding were expected to validate the predictions made from the homology structure regarding enzyme-substrate interactions.

Expression and Purification of WNV CF40.Gly.NS3pro and Mutants—The recombinant WNV protease (WNV CF40.Gly.NS3pro) used in this study comprises the essential 40 amino acid cofactor domain of NS2B joined to the 184 amino acid protease domain of NS3 by a flexible nonapeptide G4SG4,a b-amidase domain (25). Comparative kinetic studies of these mutant proteases and the assessment of the role of the mutated residues in substrate binding were expected to validate the predictions made from the homology structure regarding enzyme-substrate interactions.

Expression and Purification of WNV CF40.Gly.NS3pro and Mutants—The recombinant WNV protease (WNV CF40.Gly.NS3pro) used in this study comprises the essential 40 amino acid cofactor domain of NS2B joined to the 184 amino acid protease domain of NS3 by a flexible nonapeptide G4SG4,a b-amidase domain (25). Comparative kinetic studies of these mutant proteases and the assessment of the role of the mutated residues in substrate binding were expected to validate the predictions made from the homology structure regarding enzyme-substrate interactions.
assays were conducted in a final reaction volume of 200 μl containing 50 mM Tris, pH 9.5, 30% glycerol, and 1 mM CHAPS. Wild type, V115A, D129E, D129N, Y150F, and S160A were assayed in triplicate at an enzyme concentration of 0.5 μM and eight different Ac-KPGLKR- pNa concentrations ranging from 62.5 μM to 1 mM. D129A was assayed in quadruplicate at the higher enzyme concentration of 2.5 μM and eight Ac-KPGLKR-pNa concentrations ranging from 125 μM to 2 mM.

### Table II

|          | Km (μM) | kcat (s⁻¹) | kcat/Km (μM⁻¹ s⁻¹) |
|----------|---------|------------|---------------------|
| Wild type| 170 ± 14| 0.965 ± 0.026| 5694 ± 345          |
| V115A    | 360 ± 35| 0.572 ± 0.036| 2420 ± 141          |
| D129A    | 2942 ± 730| 0.011 ± 0.001| 3.8 ± 0.3           |
| D129E    | 561 ± 90 | 0.019 ± 0.002| 34 ± 2.6            |
| D129N    | 611 ± 56 | 0.038 ± 0.002| 62 ± 2.9            |
| Y150F    | 246 ± 17 | 0.133 ± 0.004| 541 ± 26            |
| S160A    | 357 ± 29 | 0.394 ± 0.014| 1102 ± 55           |

### Discussion

Correlation of Mutagenesis Results to Substrate Interactions—The effect of site-directed mutagenesis on protease efficiency is frequently used to provide evidence for the importance of individual enzyme residues in substrate binding and processing. Focused mutagenesis of protease residues identified from our homology structure and predicted to be important for interactions with P1 and P2 residues of the dabisac substrate proved to be quite revealing.

Tyrosine 150—Tyrosine 150 is conserved throughout flavivirus NS3 proteases. It lies at the base of the S1 pocket in the homology model of WNV NS3pro, just as observed in Dengue-2 virus NS3pro by crystallographic data (24, 29). The aromatic ring of Tyr-150 is predicted to make a primary p-cation interaction with the basic Arg residue at P1 of the substrate. Substitution of the Tyr-150 residue by alanine resulted in an inactive enzyme, consistent with previous mutagenesis results for Dengue-2 virus (38) and supporting the prediction that Tyr-150 makes a major interaction with the P1 residue.

The conservative mutation of Y150F that simply removes the hydroxyl group from the aromatic ring did not substantially alter catalytic activity but did cause a 10-fold reduction in catalytic efficiency and a small decrease in substrate affinity (1.4-fold increase in Km). This result supports the proposed p-cation interaction that requires only the aromatic ring of both of these S1 residues. The decrease in catalytic efficiency might be attributed to the removal of a hydrogen bond between the tyrosine hydroxyl group and P1 of the substrate or alternatively (and more likely) to an increase in the energy of the aromatic p-orbitals.

Serine 163—Serine 163 is also conserved throughout flavivirus NS3 proteases. The homology model predicts Ser-163 to line the S1 pocket, and this residue has been suggested to make an interaction with a P1 Arg in the crystal structure of the Bowman-Birk inhibitor bound to Dengue-2 virus NS3pro (29). Mutation of Ser-163 to alanine inactivated the protease, confirming that Ser-163 makes a critical interaction with the P1 Arg, presumably via a hydrogen bond from its hydroxyl group.

Aspartate 129—Aspartate 129 is also conserved throughout flavivirus NS3 proteases. Based on homology to trypsin, Asp-129 (Asp-189 of trypsin) was originally believed to lie at the
base of the S1 pocket and make the primary charge stabilization interaction with the substrate P1 arginine (17). However crystallization of the Dengue-2 virus NS3pro and homology to HCV NS3pro showed that Tyr-150 lies at the base of the S1 protease pocket. The extent of the substrate interaction supplied by Asp-129 was also queried in earlier cell-based mutagenesis studies, which showed that its substitution in Dengue-2 virus NS3 had a limited effect on protease activity (38). In contrast to the cell-based findings (38), kinetic analysis in this study revealed that the D129A mutation substantially reduced protease activity in vitro. Thus, Asp-129 must provide an important interaction with the substrate, presumably via a salt bridge or hydrogen bond with the positive charge of the P1 Arg. Crystallographic data for Dengue-2 virus NS3pro with Bowman-Birk inhibitor show that P1 Arg possibly exists in two conformations, one where it is bound to Asp-129 and one where it is bound to Tyr-150 and Ser-163 (29). If this is correct, Asp-129 may play a role in the initial binding of the P1 Arg before it translocates deeper into the S1 pocket where it can subsequently bind to Tyr-150 and Ser-163 during cleavage. The kinetic parameters generated for the D129A mutant are consistent with this proposal, the large $K_m$ value (17-fold larger than the wild type) suggesting that D129 is probably involved in substrate binding.

The conservative substitution D129E resulted in a 170-fold decrease in catalytic efficiency, a small decrease in substrate affinity (3-fold increase in $K_m$), and a large decrease in $k_{cat}$. Because binding affinity is only slightly decreased, the D129E increase in side chain length must interfere in the catalytic mechanism rather than retarding substrate binding. This would suggest that binding to Tyr-150 and Ser-163 without first interacting with Asp-129 positions the substrate differently, resulting in less efficient cleavage. Alternatively, if Asp-129 provides an initial interaction that is involved in the translocation of the P1 Arg into the S1 pocket, the larger side chain may not deliver the P1 Arg into its correct binding position with Tyr-150 and Ser-163 but hold it in a different position, which would result in inefficient cleavage.

The isosteric replacement D129N mutant also displayed decreased (90-fold) catalytic efficiency and a small decrease in substrate affinity (3.6-fold increase in $K_m$); however, this mutant performed better than D129A or D129E. This isosteric replacement of an acid side chain by an amide side chain seems to support the notion that a hydrogen bond interaction is important between residue 129 and the P1 Arg of substrate, but either is not as pronounced for Asn as Asp or is supplemented by a salt bridge interaction as in the case of D129.

Interestingly, whereas the kinetic analysis of the Asp-129 mutants showed a dramatic decrease in protease efficiency, the autocatalytic activity of Asp-129 mutants observed in this study appears to be close to the wild type level (Fig. 4), similar to an observation made (38) for autocatalysis by Dengue-2 virus NS2B-NS3pro D129 mutants. These findings may simply be the consequence of cis-cleavage of the proximally juxtaposed residues during protease expression. Alternatively, this comparison of different catalytic efficiencies may not be relevant because the latter case involves autocatalytic cleavage of an extended substrate, whereas the former case involves cleavage of only a six-residue substrate. Intermolecular forces outside the six-residue sequence may contribute to the association of substrate and the protease active site.

If Asp-129 does provide an initial interaction with the P1 Arg as suggested above, this may be exploited in the design of improved inhibitors. By designing an inhibitor to make simultaneous interactions with Tyr-150, Ser-163, and Asp-129, the binding affinity would be significantly greater than that of the P1 Arg and therefore should provide a very effective competitive inhibitor.

Asparagine 152—Asparagine 152 was shown by the homology model of WNV NS3pro to line the S2 pocket and is likely to make an interaction with the P2 lysine. The mutation of Asn-152 to alanine inactivated the protease, supporting the prediction that it provides an essential interaction with the P2 Lys via a hydrogen bond. Consistent with the importance of interactions with both P1 and P2 basic substrate residues, autocatalytic cleavage was not observed during purifications of any of the mutant enzymes containing Y150A, S163A, or N152A.

Valine 115—The homology model suggested that valine 115 lines the S1 pocket. Val-115 was not expected to provide a substrate interaction but was included in this mutagenesis study to investigate the size of the S1 pocket. Replacing Val-115 with the smaller alanine caused only a 2-fold decrease in catalytic efficiency because of a decrease in substrate affinity (2-fold increase in $K_m$). This may be the result of a small increase in the size of the S1 binding pocket for the occupying P1 Arg side chain, which is less than optimal. Substitution of a large phenylalanine residue for Val-115 inactivates the protease most probably because the S1 pocket becomes too small to accommodate binding of the P1 Arg.

Serine 160—Serine 160 is not highly conserved in flavivirus NS3 proteases, but the homology model of WNV NS3pro suggested that Ser-160 may contribute an interaction to the P1 side chain. The S160A mutation resulted in a 5-fold decrease in catalytic efficiency and a decrease in substrate affinity (2-fold increase in $K_m$), suggesting that Ser-160 may provide a weak interaction with the substrate, possibly via a hydrogen bond. This observation would represent a variation in substrate recognition between members of the NS3 protease family.

CONCLUSION

This study implicates some important requirements in the WNV NS2B/NS3 protease for recognition of a prototype dibasic substrate sequence. It has demonstrated that the enzyme residues at both S1 and S2 pockets are essential for substrate recognition and catalytic efficiency.

Most of the predictions made for enzyme-substrate recognition, based on our homology structure of WNV NS3pro, have been validated by comparing kinetic properties of wild type with site-specific mutant enzyme variants. The model predicted the S1 enzyme subsite to be the largest cleft in the substrate binding groove. This site was investigated in some detail in this study, and the key interactions implicated were a cation interaction between the Tyr-150 and P1 Arg side chain, a hydrogen bond between the hydroxyl group of Ser-163 and P1 Arg, a salt bridge or strong H-bond between Asp-129 and P1 Arg side chain, and a possible weak hydrogen bond supplied by Ser-160. Mutations at residue 129 support the idea that Asp-129 makes an initial interaction with the P1 Arg side chain before the latter pushes deeper into the S1 pocket where it binds with Tyr-150 and Ser-163. This could be an important feature for exploitation in the design of an inhibitor. It is conceivable that an inhibitor could be designed to interact with all three residues leading to high affinity.

Only residue Asn-152 was investigated in the S2 subsite of WNV CF40.Gly.NS3pro. However, this residue was definitively shown to be crucial for the processing of a substrate with Lys at P2, a result that implicates an important hydrogen bond between Asn-152 and the P2 Lys. Further mutagenesis could be targeted to other residues lining the S2 pocket to search for additional interactions.

We conclude that residues in both the S1 and S2 subsites of a recombinant WNV NS3 protease with tethered cofactor were crucial for substrate processing. These results immediately
suggest the prospect of designing inhibitors that target both protease subtypes. The dibasic P2/P1 requirement for substrate recognition is unusual in human serine proteases. Indeed, we have already shown that conventional serine protease inhibitors do not inhibit WNV CP40.Gly.NS3pro (25). Host cells do contain some enzymes that process proteins at pairs of basic residues with often a requirement of a basic residue at P4 or P6. Proprotein convertases like furin, for example, recognize paired basic residues and sometimes have a P6 basic residue (4). However, basic residues at P4 or P6 are not required (or desirable) at P4 or P6 in substrates for processing by WNV NS3 protease. The S2/S1 subsites may therefore represent ideal targets for design of serine protease inhibitors that are selective for this and possibly other flavivirus serine proteases over human host serine proteases.

REFERENCES

1. Hayes, C. G. (2001) Ann. N. Y. Acad. Sci. 951, 25–37
2. Brinton, M. A. (2002) Annu. Rev. Microbiol. 56, 371–402
3. Lancistti, R. S., Roehrig, J. T., Deubel, V., Smith, J., Parker, M., Steele, K., Crise, B., Volpe, K. E., Crabtree, M. B., Scherret, J. H., Hall, R. A., Mackenzie, J. S., Crupp, C. B., Panigrahy, B., Ostlund, E., Schmitt, B., Malkinson, M., Banet, C., Weissman, J., Komar, N., Savage, H. M., Stone, W., McNamara, T., and Gubler, D. J. (1999) Science 286, 2333–2337
4. Henrich, S., Cameron, A., Bourenkov G. P., Kiefersauer R., Huber, R., Lindberg, I., Bode, W., and Than, M. E. (2003) Nat. Struct. Biol. 10, 520–526
5. Buckley, A., Dawson, A., Moss, S. R., Hinsley, S. A., Bellamy, P. E., and Gould, I. Bode, W., and Than, M. E. (2003) J. Gen. Virol. 84, 2807–2817
6. Chambers, T. J., Hahn, C. S., Galler, R., and Rice, C. M. (1990) Proc. Natl. Acad. Sci. U. S. A. 87, 8898–8902
7. Chambres, T. J., Hahn, C. S., Galler, R., and Rice, C. M. (1990) Annu. Rev. Microbiol. 44, 649–688
8. Chambers, T. J., Weir, R. C., Grakoui, A., McCourt, D. W., Bazan, J. F., Fletterick, R. J., and Rice, C. M. (1990) Proc. Natl. Acad. Sci. U. S. A. 87, 8898–8902
9. Lobigs, M. (1993) Proc. Natl. Acad. Sci. U. S. A. 90, 6218–6222
10. Amberg, S. M., and Rice, C. M. (1999) J. Virol. 73, 8083–8094
11. Lin, C., Chambers, T. J., and Rice, C. M. (1993) J. Biol. Chem. 268, 20310–20316
12. Nestorowicz, A., Chambers, T. J., and Rice, C. M. (1994) J. Biol. Chem. 269, 114–123
13. Chambers, T. J., Nestorowicz, A., and Rice, C. M. (1995) J. Virol. 69, 1600–1605
14. Gerbalenya, A. E., Donchenko, A. P., Koonin, E. V., and Blinov, V. M. (1989) Nucleic Acids Res. 17, 3889–3897
15. Wengler, G. (1991) Virology 184, 707–715
16. Li, H., Clum, S., You, S., Ehner, K. E., and Padmanabhan, R. (1999) J. Virol. 73, 3108–3116
17. Bazan, J. F., and Fletterick, R. J. (1989) Virology 171, 637–639
18. Falgeut, B., Pethel, M., Zhang, Y. M., and Lai, C. J. (1991) J. Virol. 65, 467–475
19. Leung, D., Schroder, K., White, H., Fang, N. X., Stoermer, M. J., Abbenante, G., Martin, J. L., Young, P. R., and Fairlie, D. P. (2001) J. Biol. Chem. 276, 45762–45771
20. Falgeut, B., Miller, R. H., and Lai, C. J. (1993) J. Virol. 67, 2034–2042
21. Chambers, T. J., Hahn, C. S., Galler, R., and Rice, C. M. (1993) Proc. Natl. Acad. Sci. U. S. A. 90, 1167–1177
22. Love, R. A., Parge, H. E., Wickersham, J. A., Hostos, Z., Habuka, N., Moowaw, E. W., Adachi, T., Margosiak, S., Dagostino, E., and Hostoska, Z. (1998) Clin. Diagn. Virol. 10, 151–156
23. Murphy, H. M., Chambers, S. P., Markland, W., Lepre, C. A., O’Malley, E. T., Harbeson, S. L., Rice, C. M., Murcko, M. A., Caron, P. R., and Thomson, J. A. (1996) Cell 87, 343–355
24. Murphy, H. M., Judge, K., Delucas, L., and Padmanabhan, R. (2000) J. Mol. Biol. 301, 759–767
25. Graf, L., Craik, C. S., Patthy, A., Roczniai, S., Fletterick, R. J., and Rutter, W. J. (1987) Biochemistry 26, 2616–2623
26. Gussio, R., Pattabiraman, N., Zaharevitz, D. W., Kellogg, G. E., Topol, I. A., Rice, W. G., Schaeffer, C. A., Erickson, J. W., and Burt, S. K. (1996) J. Med. Chem. 39, 1645–1650
27. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) Nucleic Acids Res. 22, 4673–4680
28. Jones, D. T. (1999) J. Mol. Biol. 292, 195–202
29. Hicks, R. P., Mones, E., Kim, H., Kiser, B. W., Nichols, D. A., and Bhattacharyya, A. K. (2003) Biopolymers 68, 459–470
30. Jones, G., Willett, P., and Glen, R. C. (1995) J. Mol. Biol. 245, 43–53
31. Fairlie, D. P., Tyndall, J. D. A., Reid, R. C., Wong, A. K., Abbenante, G., Scanlon, M. J., March, D. R., Bergman, D. A., Chai, C. L. L., and Burkett, B. A. (2000) J. Med. Chem. 43, 1271–1281
32. Leung, D., Abbenante, G., Fairlie, D. P. (2000) J. Med. Chem. 43, 305–341
33. Valle, R. P., and Falgeut, B. (1998) J. Virol. 72, 624–632
34. Abbenante, G., Leung, D., Bond, T., and Fairlie, D. P. (2001) Lett. Pept. Sci. 7, 347–351
