THE JOURNAL OF BIOLOGICAL CHEMISTRY Vol. 279, No. 47, Issue of November 19, pp. 48535–48542, 2004

Printed in U.S.A.

Enzymatic Characterization and Homology Model of a Catalytically Active Recombinant West Nile Virus NS3 Protease*

Received for publication, June 18, 2004, and in revised form, August 9, 2004
Published, JBC Papers in Press, August 18, 2004, DOI 10.1074/jbc.M406810200

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

From the ‡Centre for Drug Design and Development, Institute for Molecular Bioscience, and the §Department of Microbiology and Parasitology, School of Molecular and Microbial Sciences, University of Queensland, Brisbane, Queensland 4072, Australia

West Nile Virus (WNV) is a mosquito-borne flavivirus with a rapidly expanding global distribution. Infection causes severe neurological disease and fatalities in both human and animal hosts. The West Nile viral protease (NS2B-NS3) is essential for post-translational processing in host-infected cells of a viral polypeptide precursor into structural and functional viral proteins, and its inhibition could represent a potential treatment for viral infections. This article describes the design, expression, and enzymatic characterization of a catalytically active recombinant WNV protease, consisting of a 40-residue component of cofactor NS2B tethered via a non-cleavable nonapeptide (G_4SG_4) to the N-terminal 184 residues of NS3. A chromogenic assay using synthetic *para*-nitroanilide (*p*NA) hexapeptide substrates was used to identify optimal enzyme-processing conditions (pH 9.5, $I < 0.1$ M, 30% glycerol, 1 mM CHAPS), preferred substrate cleavage sites, and the first competitive inhibitor (Ac-FASGKR-H, IC_{50} ~1 μM). A putative three-dimensional structure of WNV protease, created through homology modeling based on the crystal structures of Dengue-2 and Hepatitis C NS3 viral proteases, provides some valuable insights for structure-based design of potent and selective inhibitors of WNV protease.

The West Nile Virus (WNV)[1] was first detected in a woman living in the West Nile region of Uganda almost 70 years ago (1). WNV is a member of the *Flaviviridae* family that also includes hepatitis C, Kunjin, yellow fever, Dengue, St. Louis, and Japanese encephalitis viruses (1). WNV is transmitted via mosquitoes from avian reservoir hosts to vertebrate dead-end hosts that include humans and horses. Endemic in humans in parts of Africa, Europe, and the Middle East, previous viral outbreaks were mostly detected as asymptomatic infections or mild febrile disease states (2). However, recent outbreaks in Israel (1998), Romania (1996), and the United States (1999) (3) have been associated with more serious neurological pathology and fatal infections. During the last 5 years, WNV 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),[2] Canada, and Mexico, as well as appearing recently in the United Kingdom (5). This rapid global spread, even through developed countries, has prompted widespread implementation of prevention strategies. No vaccines or therapeutic treatments for WNV infections are yet available.

Like other members of the *Flaviviridae* family, WNV contains a single-stranded, positive-sense RNA genome, which encodes three structural proteins (capsid (C), membrane (M), envelope (E)), and seven non-structural proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B, NS5) (1, 3). The genome is initially translated as a single, large polyprotein precursor. Post-translational cleavages of the viral polyprotein precursor in the cytoplasm of host-infected cells are necessary to produce the separate structural and functional viral proteins that are essential for assembly of new viral progeny. These cleavages are performed by both host enzymes (including signalases and furin) and by a viral protease encoded by the N-terminal-third of NS3 (Fig. 1). This NS3 viral protease, which is a trypsin-like serine protease with a functional catalytic triad (His^{51}, Asp^{75}, Ser^{135}), is absolutely essential (along with viral-encoded cofactor NS2B) for viral replication and is itself autocatalytically cleaved from the polypeptide precursor. By analogy with the homologous Dengue virus NS3 protease, for which our truncation studies had previously shown that a 40-residue hydrophilic domain of NS2B is necessary for catalytic activity of the NS3 protease, (6) we infer below that a similar sequence within the WNV NS2B is likely necessary for catalytic activity of the WNV NS3 protease. The remainder of NS2B encodes 3 hydrophobic domains, believed to be involved in membrane association and localization of the enzyme to convoluted membrane regions on the endoplasmic reticulum (7). The structure of WNV NS3 protease is unknown, but there are reported crystal structures for related NS3 proteases of hepatitis C virus (HCV) (8) and Dengue virus (DEN2) (9). The structure of the latter protease was however determined in the absence of the NS2B cofactor and was thus an inactive enzyme.

By virtue of its essential function in post-translational processing of the viral polypeptide, WNV NS3 protease is implicated as a potential target for development of therapeutics that

* 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.

¶ Present address: National School of Pharmacy, University of Otago, P.O. Box 913 Dunedin, New Zealand.

‖ To whom correspondence should be addressed. Fax: 61-733462101; E-mail: d.fairlie@imb.uq.edu.au.

[1] The abbreviations used are: WNV, West Nile Virus; CAPS, 3-(cyclohexylamino)propanesulfonic acid; CHAPS, 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate; DCM, dichloromethane; DIPEA, *N*,*N*-diisopropylethylamine; DMF, *N*,*N*-dimethylformamide; ESMS, electrospray mass spectrometry; HCV, hepatitis C virus; HRMS, High Resolution electrospray mass spectrometry; IPTG, isopropyl-D-thiogalactopyranose; MES, 2-(*N*-morpholino)ethanesulfonic acid; MOPS, 4-morpholinepropanesulfonic acid; PMSF, phenylmethylsulfonyl fluoride; R_t, retention time; TLCK, tosyl-L-lysine chloromethyl ketone; HIV, human immunodeficiency virus; HPLC, high performance liquid chromatography; *p*NA, *para*-nitroanilide.

[2] CDC (2000) Division of Vector-Borne Infectious Disease: West Nile Virus; Centre for Disease Control and Prevention, Atlanta, GA, www.cdc.gov/ncidod/dvbid/westnile/index.htm.

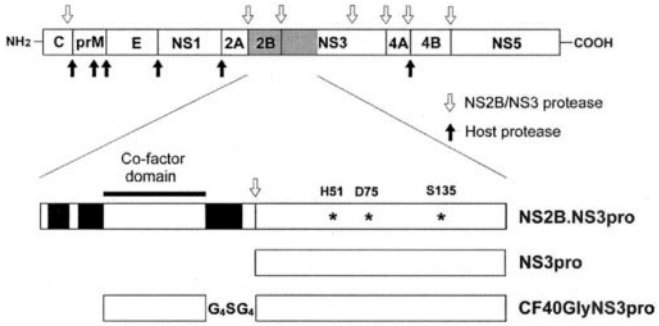

FIG. 1. **Native processing of West Nile Virus polyprotein by proteases.** Polyprotein processing by host cell proteases (*upper scheme*, *black arrows*) and by virus-encoded NS2B/NS3 protease (*open arrows*). The proteolytic domain of NS3 (*NS3pro*) and NS2B cofactor (*shaded*) is expanded (*lower scheme*) to show cleavage site (*open arrow*), catalytic triad (*asterisks*), hydrophobic domains of NS2B (*solid boxes*), and minimum 40-residue domain of cofactor NS2B required for catalytic activity. The recombinant constructs used herein: NS3pro consists of the NS3pro domain in the absence of cofactor and CF40.Gly.NS3pro comprises this 40-residue cofactor domain (*CF40*) linked via a flexible sequence (GGGGSGGGG) to NS3pro.

prevent viral replication. Precedents for inhibitors of viral proteases being potential antiviral drugs lie in the success of inhibitors of proteases of the human immunodeficiency viruses (HIV), currently the most effective treatments for humans with HIV/AIDS (10, 11), and in the early promise being shown by inhibitors of the NS3 protease of hepatitis C virus (12). This article describes the design and construction of a catalytically active recombinant WNV NS2B-NS3 protease enzyme construct, the characterization of factors that influence enzymatic processing of short synthetic chromogenic substrates containing *para*-nitroanilide (*p*NA), the first reported inhibitor of WNV protease, and the development of a three-dimensional homology-modeled structure of the WNV NS3 protease based on the known crystal structures of Dengue NS3 protease and HCV NS3 protease complexed with cofactor NS4A. It is expected that the recombinant protease cofactor construct, chromogenic assay, and homology structure might be useful for future structure-based design of even more potent and selective inhibitors of the WNV NS3 protease. Such inhibitors might form the basis for developing an antiviral therapy for the treatment of WNV infections.

EXPERIMENTAL PROCEDURES

General Methods—Protected amino acids and resins were obtained from Auspep, Novabiochem and PepChem. Trifluoroacetic acid, piperidine, DIPEA, DCM, and DMF (peptide synthesis grade) were purchased from Auspep. All other materials were reagent grade unless otherwise stated. Crude peptides were purified by reverse-phase HPLC on a Vydac C18, 10–15 μm, 300 Å, 50 × 250 mm) column, using a gradient mixture of solvent (A) 0.1% trifluoroacetic acid/water and (B) 0.1% trifluoroacetic acid/10% water/90% acetonitrile. Analytical HPLC was performed on a Waters system equipped with a 717 plus autosampler, 660 controller and a 996 photodiode array detector, using a reverse-phase Phenomenex Luna C18, 5 μm, 100 Å, 250 × 4.6-mm column. Purified peptides were characterized by analytical HPLC (linear gradient 0–100% B over 30 min), mass spectrometry, and NMR. The molecular weight of the peptides (1 mg/ml) was determined by electrospray mass spectrometry on a Micromass LCT mass spectrometer. ^{1}H NMR spectra were recorded on samples containing 4 mM peptide in Me_2SO-d_6 (550 μl) on a Bruker Avance 600 spectrometer at 298 K. Proton assignments were determined by TOCSY (80 ms mixing time), DQF-COSY, ECOSY, and NOESY (350 ms mixing time) spectra using the sequential assignment method (13). All spectra were processed on Silicon Graphics R10000 or R12000 workstations using XWINNMR 2.6 (Bruker BioSpin GmbH, Rheinstetten, Federal Republic of Germany).

Plasmid Construction—WNV (strain NY99–4132)-infected Vero cell RNA was generously provided by R. A. Hall and D. J. Nisbet (University of Queensland, Australia). cDNA was generated using the WNV NS3pro.RHindIII reverse primer (5′-GCCCCCAAGCTTACAGCATCTCAGGTTCGAAT-3′; restriction sites are underlined), which was then used as template in the generation of the WNV NS3pro and the WNV CF40.Gly.NS3pro constructs. WNV NS3pro was generated by a PCR reaction using the primers WNV NS3pro.FBamH1 5′-GCCCCTGGATCCGGTGGCGTATTGTGGGATACT-3′ and WNV NS3pro.RHindIII 5′-GCCCCCAAGCTTACAGCATCTCAGGTTCGAAT-3′. WNV CF40.GLY.NS3PRO was generated in two separate PCR reactions employing the primer pairs WNV CF40.FBamH1 5′-CGATGAC GGCGGATCCACAGATATGTGGATTGAGAGAACG-3′/WNV CF40.Gly.R 5′-CCCGCCTCCACCACTACCTCCCCCGCCCTTCCAAGGAGCACCTGGATC-3′ and WNV NS3proGly.F 5′-GGCGGGGGAGGTAGTGGTGGAGGCGGGGGAGGCGTATTGTGGGA TACT-3′/WNV NS3pro.RHindIII 5′-GCCCCCAAGCTTACAGCATCTCAGGTTCGAAT-3′ to generate PCR products encoding the essential 40 residue cofactor domain of NS2B and the 184 residue NS3pro genes, flanked with partial glycine linker sequences. These PCR products were then joined by overlapping PCR using the outer primers WNV CF40.F*Bam*HI and WNV NS3pro.RHindIII. The resulting PCR products, designated WNV NS3pro and CF40.Gly.NS3pro, comprised the 184 residue NS3 protease domain alone and the 40 residue essential cofactor domain of NS2B linked to the 184 residue NS3 protease via a flexible nonapeptide glycine linker ($Gly_4SerGly_4$), respectively. These were then digested with BamHI and HindIII and cloned into pQE9. The identity of these constructs, designated pQE9 WNV NS3pro and pQE9 WNV CF40.Gly.NS3pro was confirmed by automated sequence analysis.

Enzyme Expression and Purification—The pQE9 WNV CF40.Gly.NS3pro vector was used to allow high level, inducible expression of N-terminal hexahistidine-tagged recombinant proteins. Cultures of *Escherichia coli* strain SG13009 transformed with the expression plasmids were grown in 500 ml of LB medium containing 100 μg/ml ampicillin and 25 μg/ml kanamycin at 37 °C until OD_{600} reached 0.5. Expression of the recombinant protein was induced by the addition of isopropyl-β-D-thiogalactopyranose (IPTG) to a final concentration of 0.3 mM and incubated for an additional 3 h at 22 °C. The cells were then harvested by centrifugation and stored at −20 °C.

For protein purification, the cell pellets were thawed and resuspended in 5 ml of lysis buffer (50 mM HEPES, pH 7.5, 300 mM NaCl,10 mM imidazole, 5% glycerol) per 1 g of wet pellet. To prevent proteolytic cleavage of protein during cell lysis and purification the following protease inhibitors were added to give the final concentration of 1 μg/ml aprotinin, 1 μg/ml leupeptin, 1 μg/ml benzamide, 1 mM PMSF, and 10 μg/ml TLCK. The resuspended cells were lysed by three passes through a French Press at a pressure of 1000 Pascals and then centrifuged at 27,000 × *g* for 20 min at 4 °C. The recombinant proteases, tagged at the N terminus by a hexahistidine sequence, were purified by affinity chromatography on a Ni^{2+}-nitrilotriacetic acid-agarose column (Qiagen). A 1-cm column was pre-equilibrated with 30 ml of column buffer (50 mM HEPES, pH 7.5, 300 mM NaCl, 10 mM imidazole, 5% glycerol), and then the resin was removed and mixed with the supernatant of the cell lysates. These mixtures were then incubated at 4 °C on a shaker for 30 min to allow the His-tagged protein to bind to the Ni^{2+} resin. The columns were re-packed and washed with 30 ml of column buffer containing 20 mM imidazole before eluting the protein into 6 × 1 ml fractions with column buffer containing 100 mM imidazole. The pre- and post-IPTG induction samples, soluble and insoluble fractions, and the elution fractions were analyzed by 12% SDS-PAGE.

Substrate Synthesis—The *p*NA substrates corresponding to the native WNV NS3 protease cleavage sites Ac-DPNRKR-*p*NA (NS2A-NS2B), Ac-LQYTKR-*p*NA (NS2B-NS3), Ac-FASGKR-*p*NA (NS3-NS4A), and Ac-KPGLKR-*p*NA (NS4A-NS5) were synthesized by a reported method (15).

Ac-DPNRKR-pNA—R_t = 16.4 min. ESMS [M+H] = 947.8, [M+2H]/2 = 474.4. ^{1}H NMR (Me_2SO-d_6) δ 12.37 (br s, 1H, AspδOH), 10.58 (s, 1H, *pNA*-NH), 8.27 (d, 1H, AspαNH), 8.22 (d, 2H, *pNA*ArH), 8.15 (d, 1H, Arg6αNH), 8.00 (d, 1H, AsnαNH), 7.86 (d, 1H, LysαNH),7.84 (d, 2H, *pNA*-ArH), 7.73 (d, 1H, Arg4αNH), 7.67 (m, 3H, LyszNH_3), 7.60 (t, 1H, Arg6ϵNH),7.50 (t, 1H, Arg4ϵNH), 7.5–6.6 (8H, m, 4xArgNH_2), 7.42 (s, 1H, AsnδNH), 6.97 (s, 1H, AsnδNH), 4.81 (m, 1H, AspαCH), 4.45 (m, 1H, AsnαCH), 4.34 (m, 1H, Arg6αCH), 4.25 (dd, 1H, ProαCH), 4.22 (m, 1H, LysαCH), 4.21 (m, 1H, Arg4αCH), 3.68 (m, 1H, ProδCH),3.61 (m, 1H, ProδCH), 3.12 (m, 2H, Arg6$\delta$$CH_2$), 3.07 (m, 2H, Arg4$\delta$$CH_2$), 2.73 (m, 2H, Lys$\epsilon$$CH_2$), 2.68 (dd, 1H, Asp$\beta$CH), 2.58 (dd, 1H, Asn$\beta$CH), 2.44 (m, 1H, Asp$\beta$CH), 2.43 (m, 2H, Asn$\beta$CH), 2.03 (m, 1H, Pro$\beta$CH), 1.87 (m, 2H, Pro$\gamma$$CH_2$), 1.83 (m, 1H, Pro$\beta$CH), 1.80 (s, 3H, Ac-$CH_3$), 1.77 (m, 1H, Arg4$\beta$CH), 1.77 (m, 1H, Arg6$\beta$CH), 1.67 (m, 1H, Lys$\beta$CH), 1.65 (m, 1H, Arg6$\beta$CH), 1.57 (m, 1H, Lys$\beta$CH), 1.55 (m, 1H, Arg6$\gamma$CH), 1.50 (m, 1H, Arg4$\beta$CH), 1.50 (m, 1H, Arg4$\gamma$CH), 1.50 (m, 1H, Lys$\gamma$CH), 1.50 (m, 2H,

LysδCH_2), 1.48 (m, 1H, Arg6γCH), 1.33 (m, 1H, LysγCH). HRMS calculated for $C_{39}H_{62}N_{16}O_{12}$ $[M+H]^+$ = 947.4806, found 947.4782; [M+2H]/2 = 474.2439, found 474.2443.

Ac-LQYTKR-pNA—R_t = 18.7 min. ESMS [M+H] = 970.6, [M+2H]/2 = 485.8. ^{1}H NMR (Me_2SO-d_6) δ 10.60 (s, 1H, *pNA*-NH), 9.15 (s, 1H, TyrzOH), 8.26 (d, 1H, ArgαNH), 8.22 (d, 2H, *pNA*-ArH), 8.19 (d, 1H, GlnαNH), 8.07 (d, 1H, LeuαNH), 7.90 (d, 1H, LysαNH), 7.86 (d, 1H, TyrαNH), 7.84 (d, 2H, *pNA*-ArH), 7.80 (d, 1H, ThrαNH), 7.64 (m, 3H, LyszNH_3), 7.51 (t, 1H, Arg6ϵNH), 7.4–6.6 (4H, m, 2xArgNH_2), 7.24 (s, 1H, GlnϵNH), 7.00 (d, 2H, 2xTyrδCH), 6.79 (t, 1H, GlnϵNH), 6.79 (d, 2H, 2xTyrϵCH), 4.93 (s, 1H, ThrγOH), 4.46 (m, 1H, TyrαCH), 4.35 (m, 1H, ArgαCH), 4.30 (m, 1H, LysαCH), 4.22 (m, 1H,ThrαCH), 4.19 (m, 1H, LeuαCH), 4.12 (m, 1H, GlnαCH), 3.98 (m, 1H, ThrβCH), 3.11 (m, 2H, ArgδCH_2), 2.94 (dd, 1H, TyrβCH), 2.74 (m, 2H, LysϵCH_2), 2.71 (m, 1H, TyrβCH), 2.06 (m, 1H, GlnγCH), 2.00 (m, 1H, GlnγCH), 1.86 (s, 3H, Ac-CH_3), 1.80 (m, 1H, GlnβCH), 1.76 (m, 1H, ArgβCH), 1.69 (m, 1H, LysβCH), 1.69 (m, 1H, GlnβCH), 1.65 (m, 1H, ArgβCH), 1.58 (m, 1H, LeuγCH), 1.56 (m, 1H, LysβCH), 1.56 (m, 1H, ArgγCH), 1.53 (m, 1H, LysγCH), 1.51 (m, 2H, LysδCH_2), 1.48 (m, 1H, ArgγCH), 1.42 (m, 1H, LeuβCH), 1.39 (m, 1H, LeuβCH), 1.33 (m, 1H, LysγCH), 1.05 (d, 3H, ThrγCH_3), 0.86 (d, 3H, LeuδCH_3), 0.81 (d, 3H, LeuδCH_3. HRMS calculated for $C_{44}H_{67}N_{13}O_{12}$ $[M+H]^+$ = 970.5105, found 970.5098; [M+2H]/2 = 485.7589, found 485.7601.

Ac-FASGKR-pNA—R_t = 18.1 min. ESMS [M+H] = 827.6, [M+2H]/2 = 414.3. ^{1}H NMR (Me_2SO-d_6) δ 10.61 (s, 1H, *pNA*-NH), 8.30 (d, 1H, ArgαNH), 8.22 (d, 2H, *pNA*-ArH), 8.21 (d, 1H, AlaαNH), 8.12 (t, 1H, GlyαNH), 8.11 (d, 1H, PheαNH), 7.94 (d, 1H, LysαNH), 7.87 (d, 1H, SerαNH), 7.84 (d, 2H, *pNA*-ArH), 7.66 (m, 3H, LyszNH_3), 7.58 (t, 1H, ArgϵNH), 7.46–6.65 (4H, m, 2xArgNH2), 7.27–7.22 (m, 4H, PheArH), 7.20–7.14 (m, 1H, PheArH), 5.04 (s, 1H, SerOH), 4.48 (m, 1H, PheαCH), 4.30 (m, 1H, ArgαCH), 4.31 (m, 1H, AlaαCH), 4.30 (m, 1H, LysαCH), 4.23 (m, 1H, SerαCH), 3.78 (dd, 1H, GlyαCH), 3.71 (dd, 1H, GlyαCH),3.62 (m, 1H, SerβCH), 3.55 (m, 1H, SerβCH), 3.11 (m, 2H, ArgδCH_2), 2.99 (dd, 1H, PheβCH), 2.72 (m, 2H, LysϵCH_2), 2.71 (m, 1H, PheβCH), 1.76 (m, 1H, ArgβCH), 1.74 (s, 3H, Ac-CH_3), 1.66 (m, 1H, ArgβCH), 1.64 (m, 1H, LysβCH),1.56 (m, 1H, ArgγCH),1.51 (m, 1H, LysβCH), 1.50 (m, 2H, LysδCH2), 1.49 (m, 1H, LysγCH),1.48 (m, 1H, ArgγCH), 1.30 (m, 1H, LysγCH),1.22 (d, 3H, AlaβCH_3). HRMS calculated for $C_{37}H_{54}N_{12}O_{10}$ $[M+H]^+$ = 827.4159, found 827.4186; [M+2H]/2 = 414.2116, found 414.2130.

Ac-KPGLKR-pNA—R_t = 16.9 min. ESMS [M+H] = 860.7, [M+2H]/2 = 430.8. ^{1}H NMR (Me_2SO-d_6) δ 10.58 (s, 1H, *pNA*-NH), 8.23 (d, 2H, *pNA*-ArH), 8.19 (t, 1H, GlyαNH), 8.18 (d, 1H, ArgαNH), 8.05 (d, 1H, Lys1αNH), 7.95 (d, 1H, Lys5αNH), 7.84 (d, 2H, *pNA*-ArH), 7.76 (d, 1H, LeuαNH), 7.68 (m, 6H, 2xLyszNH_3), 7.7–6.5 (4H, m, 2xArgNH_2), 7.61 (t, 1H, Arg6ϵNH), 4.47 (m, 1H, Lys1αCH), 4.36 (m, 1H, ArgαCH), 4.30 (m, 1H, LeuαCH), 4.26 (m, 1H, ProαCH), 4.24 (m, 1H, Lys5αCH), 3.68 (m, 2H, GlyαCH_2), 3.67 (m, 1H, ProδCH), 3.56 (m, 1H, ProδCH), 3.11 (m, 2H, ArgδCH_2), 2.75 (m, 2H, Lys1ϵCH_2), 2.73 (m, 2H, Lys5ϵCH_2), 2.05 (m, 1H, ProβCH), 1.93 (m, 1H, ProγCH), 1.85 (m, 1H, ProγCH), 1.82 (s, 3H, Ac-CH_3), 1.80 (m, 1H, ProβCH), 1.75 (m, 1H, ArgβCH), 1.66 (m, 1H, Lys5βCH), 1.65 (m, 1H, ArgβCH), 1.63 (m, 1H, Lys1βCH), 1.56 (m, 1H, Lys5βCH), 1.56 (m, 1H, ArgγCH), 1.56 (m, 1H, LeuγCH), 1.53 (m, 1H, LysγCH), 1.51 (m, 2H, Lys1δCH_2), 1.50 (m, 2H, Lys5δCH_2), 1.49 (m, 1H, Lys1βCH), 1.49 (m, 1H, Lys1γCH), 1.48 (m, 1H, LeuβCH), 1.47 (m, 1H, ArgγCH), 1.43 (m, 1H, LeuβCH), 1.32 (m, 1H, Lys1γCH), 1.32 (m, 1H, Lys5γCH), 0.87 (d, 3H, LeuδCH_3), 0.83 (d, 3H, LeuδCH_3). HRMS calculated for $C_{39}H_{65}N_{13}O_9$ $[M+H]^+$ = 860.5101, found 860.5061; [M+2H]/2 = 430.7587, found 430.7617.

Inhibitor Synthesis—The peptide aldehyde Ac-FASGKR-H was synthesized by the method of Siev and Semple (16). Ac-FASGKR-H (R_t = 15.5 min): ESMS [M+H] = 691.4. HRMS calculated for $C_{31}H_{50}N_{10}O_8$ $[M+H]^+$ = 691.3886, found 691.3896.

Enzymatic Characterization—The recombinant WNV protease (WNV CF40.Gly.NS3pro) was assayed against four hexapeptide substrates corresponding to P6-P1 of known endogenous cleavage sites NS2A-NS2B (Ac-DPNRKR-*p*NA), NS2B-NS3 (Ac-LQYTKR-*p*NA), NS3-NS4A (Ac-FASGKR-*p*NA), NS4B-NS5 (Ac-KPGLKR-*p*NA) with a chromogenic *p*NA group in the P1′ position. Cleavage of the *p*NA chromophore from the peptide substrates by the WNV protease produced a detectable color change at 405 nm, allowing protease activity to be measured. The assay was conducted on a 96-well plate with a final reaction volume of 200 μl containing buffer (plus glycerol and CHAPS detergent for optimal processing), 0.5 μM recombinant protease, and 500 μM substrate at 37 °C. After preincubation in separate wells (10 min, 37 °C), catalysis was initiated by mixing substrate with the enzyme buffer solution by automatic shaking. The optical density was measured at 405 nm every 30 s for 10 min in a SpectraMax 250 reader, and the average change in mOD/min was calculated.

pH dependence was determined over the range 5.5–12.0 using 0.5 μM enzyme, 500 μM substrate (Ac-FASGKR-*p*NA), incubation temperature 37 °C, constant ionic strength (I = 5 mM), final volume 200 μl, for 50 mM buffer (MES (pH 5.5–7.0), MOPS (pH 6.5–8.0), Tris (pH 7.5–9.5), CAPS (pH 10.0–11.0)) containing NaCl, glycerol, CHAPS detergent.

Glycerol dependence was determined by assaying environments of increasing glycerol content, from 0 to 50% glycerol (5% increments) in the final volume of 200 μl. Glycerol dependence assays contained buffer with NaCl, glycerol, CHAPS detergent, 0.5 μM enzyme, 500 μM substrate (Ac-FASGKR-*p*NA), constant ionic strength (I = 5 mM, NaCl), at 37 °C in a final volume 200 μl.

Ionic strength dependence was determined for 0–200 mM NaCl in buffer (with glycerol, CHAPS), 0.5 μM enzyme, 500 μM substrate (Ac-FASGKR-*p*NA), 37 °C in the final volume of 200 μl.

Initial reaction velocities were monitored continuously at 405 nm on a SpectraMax® 250 spectrophotometer, and then interpreted using the accompanying SoftMax™ Pro software, both from Molecular Devices Corporation. Subsequent assays for K_m determination were performed under optimal pH, glycerol and ionic strength conditions as determined by interpretation of the previous pH, glycerol, and ionic strength assays via SoftMax™ Pro and GraphPad Prism® software.

Optimal assaying conditions, found to be 30% glycerol, pH 9.5, no added NaCl, were used to determine K_m and k_{cat} values measured for the four *p*NA-peptide substrates at concentrations ranging from 100 to 3000 μM. Results were graphed using GraphPad Prism® version 4.0a for Macintosh, GraphPad Software, San Diego, CA (www.graphpad.com), and kinetic parameters calculated from weighted nonlinear regression assuming Michaelis-Menten kinetics ($v = V_{max}[S]/[S]+K_m$), k_{cat} values calculated from $k_{cat} = V_{max}/[E]$. Data are reported as the means of triplicate measurements ± S.E.

Homology Modeling—West Nile Virus NS3 protease homology models were generated using the structures of the Dengue virus NS3 protease without ligand (pdb1BEF), (9) and with Bowman-Birk inhibitor (pdb1DF9, subunit B), (9) and hepatitis C virus NS3 protease with bound cofactor (pdb1A1Q) (8). Models were generated using the Modeler and the Homology module within the InsightII molecular modeling system (17). All modeling was performed using Silicon Graphics R10000 workstations.

Sequences were aligned using Align 2D, Structure Alignment (Homology Module, InsightII) and ClustalW alignment (18). Secondary structure predictions were conducted using web-based PsiPred prediction for each of the three NS3 protease sequences (19). Electrostatic potential mapping was performed on the WNV NS3 protease model using the Delphi module in InsightII (20). A four-residue P3-P1′ substrate (Thr-Lys-Arg-Gly) was docked into the active site of a homology model of the WNV NS3 protease using GOLD v2.1 (21). A specified distance constraint between substrate residue Arg (P1) and WNV protease Tyr^{150} (S1), as well as distance and hydrogen bond constraints between substrate residue Thr (P3) and WNV protease Gly^{153} (S3) were used based on our understanding of protease-inhibitor interactions (22, 23).[3] The docked structures were ranked for best fitting conformations using GOLD 2.1 (21).

RESULTS

Design of a Catalytically Active WNV Protease—The single serine protease within the flaviviral WNV RNA genome provides an ideal starting point and therapeutic target for research into the design and development of a WNV serine protease inhibitor. The first step in designing a catalytically active WNV protease was based on our previous work on the Dengue-2 viral protease, for which we identified a minimum sequence of 184 residues from the NS3 protease (NS3pro) and 40 residues from the NS2B cofactor (CF40) that are necessary for catalytic processing of substrates (6, 25). In that work we created a catalytically active protease construct by fusing CF40 to NS3pro separated by various linkers, the most active construct involving a nonapeptide linker (GGGGSGGGG) as a replacement for the normal NS2B-NS3A cleavage sequence.

Alignment of the WNV and Dengue virus NS2B-NS3 se-

[3] J. D. A. Tyndall, T. Nall, and D. P. Fairlie, submitted for publication.

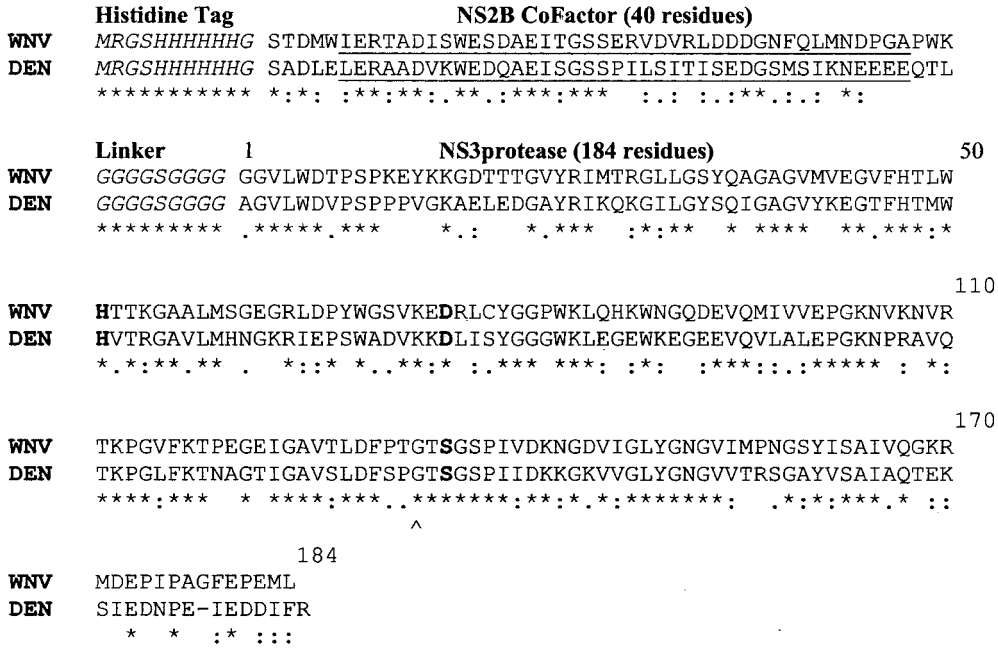

FIG. 2. **Alignment (ClustalW) of recombinant NS3-NS2B protease-cofactor sequences for Dengue (6) and West Nile Virus (18).** Alignment shows catalytic residues Ser^{135} (*S*), Asp^{75} (*D*), His^{51} (*H*) in *bold*, oxyanion residue Gly 133 (*G*). 40 residue cofactor domain of NS2B (*underlined*) as defined for Dengue type 4 virus (^) (32); exact amino acid match (⋆), high chemical similarity (:), low chemical similarity (.), insertion/deletions (−), and residues (*italics*) incorporated in construct design. Alignment is numbered from the first residue of the protease domain.

quences (Fig. 2) indicated 53% sequence identity (exact amino acid match) and 79% sequence similarity (hydrophilicity/phobicity) within the 184 residues of NS3pro. This homology, including conserved catalytic residues His^{51} (H), Asp^{75} (D), and Ser^{135} (S), and oxyanion residue Gly^{133} (G), gave us confidence in our alignment. Residue Gly^{178} in WNV NS3 was not present in Dengue-2 virus NS3. Alignment of the NS2B cofactor domains displayed 33% sequence identity (exact amino acid match) and 75% sequence similarity (high and low chemical similarity). These sequence comparisons suggested that, by analogy with the Dengure-2 NS2B-NS3 protease, a catalytically active sequence for WNV NS2B-NS3 protease might similarly constitute 40 residues from the NS2B cofactor and 184 residues from NS3, connected via a nonapeptide linker G_4SG_4 (abbreviated as Gly) in place of the normal NS2B-NS3 cleavage site sequence. We designated the putative catalytic protease construct as WNV CF40.Gly.NS3pro. A recombinant construct was also generated consisting of the 184 residue NS3 protease domain in the absence of the NS2B cofactor domain to confirm that activity was dependent on the correct association of the cofactor, as was previously shown for the Dengue virus equivalent construct (6). We designated this construct WNV NS3pro.

Expression and Purification of WNV NS3pro and WNV CF40.Gly.NS3pro—The first experimental aim was to express and purify a recombinant WNV NS3 protease domain and a catalytically active recombinant WNV protease based on the recombinant Dengue-2 virus NS3 protease (6). The homologous WNV recombinant constructs (WNV NS3pro and WNV CF40.Gly.NS3pro) were produced and cloned into the pQE9 expression vector. The sequence of these constructs and their insertion into pQE9 was confirmed by automated sequence analysis.

The pQE9 vector was used because it allowed a high level of expression, inducible with the addition of IPTG. The addition of an N-terminal hexahistidine tag enabled ready purification by Ni^{2+} affinity chromatography. The SDS-PAGE expression profile of pQE9 WNV NS3pro and pQE9 WNV CF40.Gly.NS3pro are shown in Fig. 3, *A* and *B*, respectively and the Ni^{2+} affinity chromatography of purified WNV CF40.Gly.NS3pro is shown in Fig. 3*C*. Fig. 3*A* shows a major band at ~23 kDa, present only in the bacterial lysates prepared after induction with IPTG. This band was attributed to recombinant WNV NS3pro as it is under the control of the T5 promoter and correlates approximately with the theoretical size of WNV NS3pro predicted by Vector NTI of 21 kDa. As we had observed previously with Dengue virus NS3pro (6) following extraction, the majority of the expressed product was found in the insoluble pellet fraction, with very little present in the soluble fraction. Fig. 3*B* shows a major band at ~32 kDa, present only in the bacterial lysates prepared after induction with IPTG. This band was attributed to recombinant WNV CF40.Gly.NS3pro as it is under the control of the T5 promoter and correlates approximately with the theoretical size of WNV CF40.Gly.NS3pro predicted by Vector NTI of 27 kDa. The majority of this expressed product was present in the soluble fraction, indicating that the recombinant protein was likely to be folded correctly. Both recombinants were subsequently purified from the soluble fractions by Ni^{2+} affinity chromatography. Because of the low level of recombinant WNV NS3pro in the soluble fraction relatively poor recoveries were obtained. Fig. 3*C* however shows that the WNV CF40.Gly.NS3pro recombinant (~30 kDa) was recovered and purified efficiently. The difference in molecular weight estimations between B and C was likely due to the use of prestained and non-prestained molecular weight markers, respectively. The two additional bands were attributed to products of autocatalytic cleavage, since they were not present in the purification of an inactive recombinant S135A mutant enzyme (data not shown) consistent with these fragments being dependent on proteolytic activity. These products were not expected to have an effect on processing of synthetic peptide substrates as the rate of autocatalytic cleavage is several orders of magnitude lower than that of peptide substrate cleavage and the substrate concentration is far in excess of enzyme concentration.

Proteolytic Activity—In vivo there are at least seven known post-translational cleavage sites that are processed by WNV

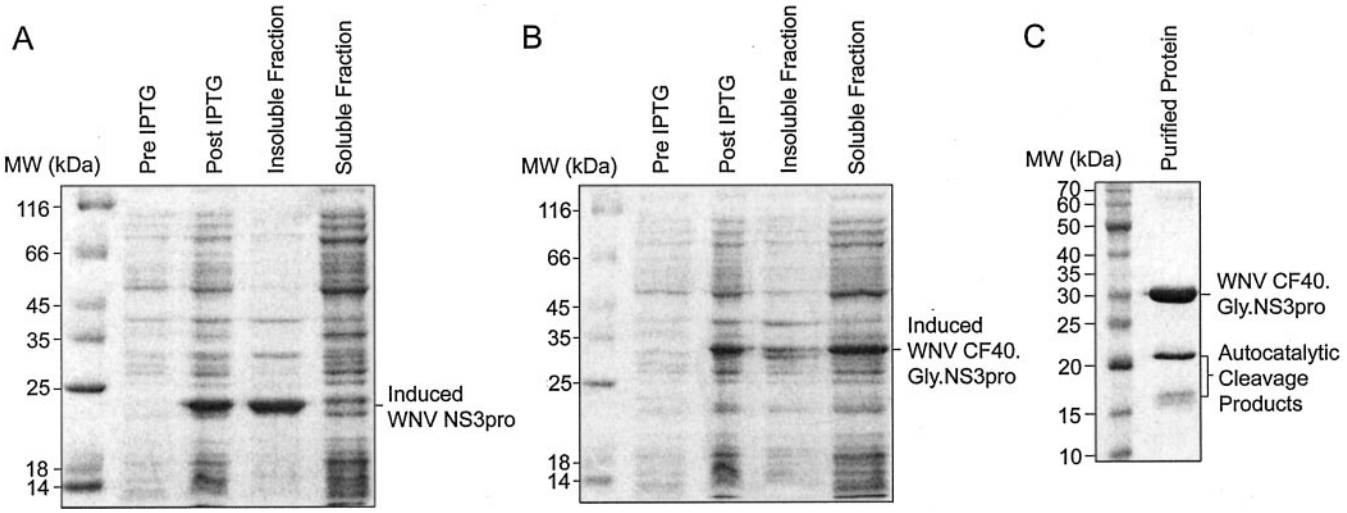

FIG. 3. **Expression of the West Nile Virus recombinant proteases.** SDS-PAGE expression profiles of WNV NS3pro (*A*) and WNV CF40.Gly.NS3pro (*B*). Lanes show cell lysate before and after the induction of recombinant protein expression with IPTG as well as the supernatant-containing soluble protein fraction and the pellet fraction containing insoluble protein after centrifugation at 13 krpm for 5 min. The *highlighted bands* are only present in samples collected after induction with IPTG and are thus suspected to be the recombinant proteins WNV NS3pro and WNV CF40.Gly.NS3pro, in *A* and *B*, respectively. *C*, SDS-PAGE purified WNV CF40.Gly.NS3pro. The major species present at approximately ~30 kDa is recombinant protease, bands at ~21 and 16 kDa are cleavage fragments.

protease in the cytoplasm of infected host cells (Fig. 1). Amino acid residues corresponding to those cleavage sites are displayed in Fig. 4. We chose P6-P1 hexapeptides from four of these sequences to construct putative substrate analogues terminating in a chromophore (*para*-nitroanilide) at P1′. The putative substrate sequences were Ac-DPNRKR-*p*NA (NS2A-2B), Ac-LQYTKR-*p*NA (NS2B-NS3), Ac-FASGKR-*p*NA (NS3-NS4A), and Ac-KPGLKR-*p*NA (NS4B-NS5). These potential substrates were to be used to characterize proteolytic activity of WNV CF40.Gly.NS3pro under a variety of assay conditions, culminating in an optimum set of conditions that could be used for more detailed kinetic analysis, and ultimately for inhibitor screening. A preliminary comparison of these four substrates suggested that Ac-FASGKR-*p*NA was processed fastest and with the highest efficiency (k_{cat}/K_m), so that particular substrate was used in the following investigation of factors that influence substrate processing by the protease. Sufficient purified WNV NS3pro was obtained to assess its enzymatic properties against Ac-FASGKR-*p*Na and was the found not to cleave the substrate, confirming that the protease is inactive in the absence of the NS2B cofactor.

Effect of pH on Enzymatic Activity—The WNV NS3 protease (WNV CF40.Gly.NS3pro) was assayed over a wide pH range from pH 5.5–11.0, using a variety of buffers suitable for each pH range: MES (pH 5.5–7.0), MOPS (pH 6.5–8.0), Tris-HCl (pH 7.5–9.5), CAPS (pH 10.0–11.0). The optimal pH for proteolytic cleavage was determined to be pH 9.5 (Fig. 5*A*). In the absence of enzyme, the substrate was resistant to hydrolysis above the optimal processing pH of 9.5, and was still intact at pH 12.0. The optimal processing pH of the WNV protease (pH 9.5) is comparable to the optimal processing pH (9.0) that we found previously for the Dengue-2 virus NS3 protease (6).

Effect of Glycerol on Enzymatic Activity—We had previously found that glycerol enhances catalytic processing by Dengue NS3 protease (6), an effect we attributed to minimizing enzyme aggregation (26). We therefore examined the effect of varying glycerol concentrations (0–50% v/v) in the final assay volume on catalytic rate of processing of FASGKR-*p*NA. There was a linear concentration dependence on glycerol (Fig. 5*B*), consistent with increasing glycerol-promoting dissociation of aggregating enzyme molecules and increasing accessibility of the enzyme active site to substrate. While increasing glycerol content of the assay environment proved beneficial to processing of substrates, the increasing viscosity of aqueous glycerol limited the assay to 30% glycerol in the final assay volume.

	$\ldots P_2 P_1 \downarrow P_1' \ldots$
C/prM	TKQK**KR**↓**G**GTAGF
NS2A-2B	DPNR**KR**↓**G**WPATE
NS2B-3	LQYT**KR**↓**G**GVLWD
NS3A-3B	SAAQ**RR**↓**G**RIGRN
NS3-4A	FASG**KR**↓**S**QIGLI
NS4A-4B	EPEK**QR**↓**S**QTDNQ
NS4B-5	KPGL**KR**↓**G**GAKGR

FIG. 4. **Native WNV polypeptide sequences cleaved by WNV protease.** Eight recognition sequences (P6-P6′) and post-translational cleavage sites (↓) known to be processed in host-infected cells by WNV protease (4, 14, 24). *Bold residues* indicate P2, P1, and P1′residues.

Effect of Ionic Strength on Enzymatic Activity—The effect of increasing ionic strength was tested by adding NaCl (0–200 mM) to the reaction medium. The results (Fig. 5*C*) clearly show that increasing ionic strength inhibits the proteolytic processing of the substrate. High NaCl might conceivably be detrimental by competing with substrate and water molecules for occupancy of the substrate binding active site, thereby inhibiting substrate processing. Alternatively, NaCl is known to favor α-helicity in peptides, and thus may alter conformational equilibria of substrates in solution, rendering them less recognizable to the protease. It has been convincingly demonstrated elsewhere that all proteases only recognize their substrates and inhibitor analogues in an extended β-strand conformation (22, 27). Thus, lower proportions of substrates in a β-strand conformation would reduce the catalytic turnover, as observed.

Enzyme Kinetics: Substrates and Inhibitor—The four peptide substrates, corresponding to the different NS2A-NS2B, NS2B-NS3, NS3-NS4A, NS4B-NS5 cleavage sites of the WNV RNA genome, were tested for their capacity to be cleaved by the WNV protease under the optimal processing conditions outlined above (50 mM ethanolamine buffer, pH 9.5; 30% glycerol; 1 mM CHAPS, no added NaCl, 37 °C, final volume 200 μl). The assay was conducted using eight different substrate concentrations ranging from 125 to 3000 μM. The results were graphed according to Michaelis-Menten kinetics $v = V_{max}[S]/([S] + K_m)$

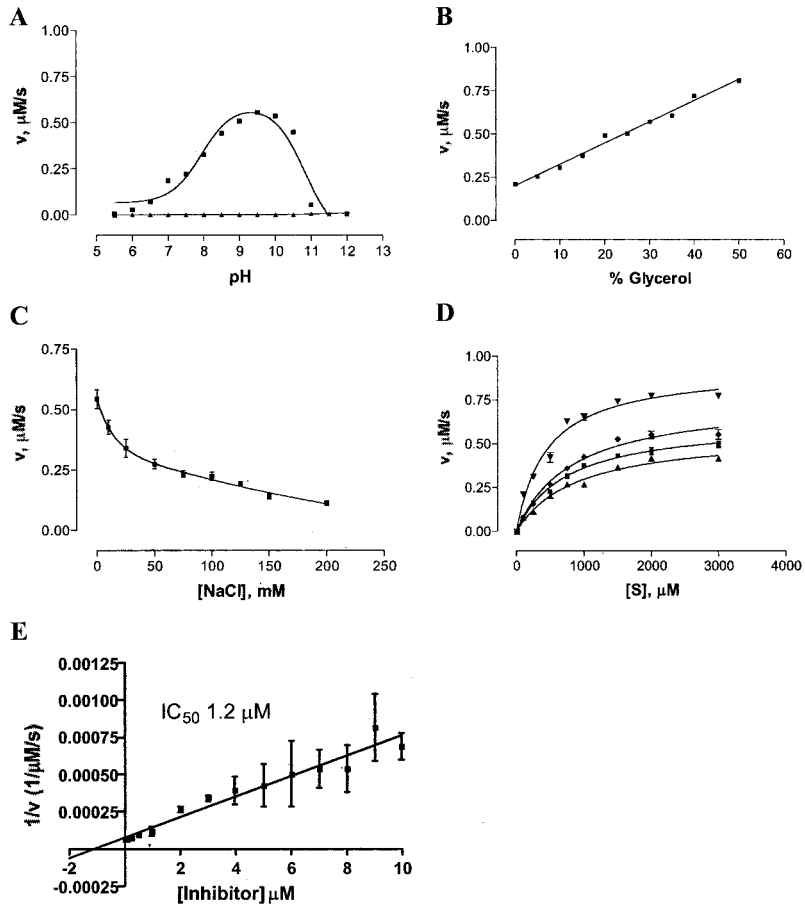

FIG. 5. **Enzymatic characterization of WNV CF40.Gly.NS3pro.** Effects of pH, buffer, glycerol content, ionic strength, and a substrate-derived inhibitor on proteolytic processing by WNV.CF40Gly.NS3pro construct. *A*, effect of pH. Substrate analogue Ac-FASGKR-*p*NA (NS3-NS4A cleavage site), ionic strength (NaCl 10 mM), glycerol content (30% of final volume) over a range of pH buffers (50 mM): MES (pH 5.5–7.0), MOPS (pH 6.5–8.0), Tris-HCl (pH 7.5–9.5), CAPS (pH 10.0–12.0) (■); and in absence of enzyme (▲). *B*, effect of glycerol. Same conditions as *A* except [glycerol] = 0–50% (■), and pH 9.0 (Tris buffer). *C*, effect of ionic strength. Same conditions as *B* except [glycerol] = 20% of final volume and varying ionic strength (NaCl: 0 mM, 10 mM, 25 mM, 50 mM, 75 mM, 100 mM, 125 mM, 150 mM, 200 mM) (■). *D*, effect of substrate on protease processing. *Para*-nitroanilide substrates corresponding to cleavage sites NS2A-2B (Ac-DPNRKR-*p*NA; ■, NS2B-NS3 (Ac-LQYTKR-*p*NA; ▲), NS3-NS4A (Ac-FASGKR-*p*NA; ▼), NS4B-NS5 (Ac-KPGLKR-*p*NA; ♦) under optimal assay conditions (50 mM ethanolamine buffer, pH 9.5, 30% glycerol, 1 mM CHAPS, 37 °C, final volume of 200 μl). *E*, effect of inhibitor. Processing of Ac-FASGKR-pNA under same conditions as in *D* but with varying concentrations (only 1–10 μM [I] shown) of aldehyde inhibitor Ac-FASGKR-H. IC_{50} ~1 μM. In all cases, *bars* indicate S.D. from the mean resulting from the average of three replicates.

(Fig. 5D) in order to calculate the binding affinity constant K_m and the reaction rate k_{cat} (= V_{max}/[E]). Kinetic data (Table I) suggested only small changes in processing rate (k_{cat}) between the four substrates, with the NS3-NS4A cleavage site being processed fastest and the NS2B-NS3 cleavage site processed slowest. The highest affinity substrate (lowest K_m) was the sequence corresponding to the NS3-NS4A cleavage site. The rank order for highest catalytic efficiency (k_{cat}/K_m) of these *p*NA substrates was AcFASGKR-*p*NA (NS3-NS4A) > Ac-KPGLKR-*p*NA (NS4B-NS5) ≥ Ac-DPNRKR-*p*NA (NS2A-NS2B) ≥ Ac-LQYTKR-*p*NA (NS2B-NS3). By comparison, corresponding processing of Dengue *p*NA substrates by Dengue-2 CF40.Gly.NS3pro (6) displayed K_m and k_{cat} values that were lower than for the WNV substrates. This suggests greater catalytic efficiency (k_{cat}/K_m) for substrate processing by the WNV NS3 protease compared with the analogous recombinant Dengue NS3 protease construct.

Based on this preference for the P6-P1 substrate sequence corresponding to NS3-NS4A, we also constructed Ac-FASGKR-H with a C-terminal aldehyde and found (Fig. 5*E*) this to be a reasonably potent competitive inhibitor (IC_{50} ~1 μM) of catalytic processing of Ac-FASGKR-*p*NA by WNV CF40.Gly.NS3pro. This potency is despite the fact that the aldehyde exists in aqueous solution as an equilibrium mixture (Fig. 6) with a cyclic form, in which the side chain guanidine of Arg has condensed with aldehyde, and with acetal formed through hydration of the aldehyde. The free aldehyde is likely to be the active form of this inhibitor (28), but is not present above 5% in the mixture in aqueous solution, the bulk of the compound existing in the hydrate form of the aldehyde and in two diastereomeric hemi-aminal cyclic forms (Fig. 6). The modest potency observed for this inhibitor may be a consequence of the unfavorable position of this equilibrium, with little of the active

TABLE I
Kinetic parameters for the hydrolysis of synthetic peptide substrates by WNVCF40.Gly.NS3pro

Hydrolysis was performed at 37 °C, pH 9.5, ethanolamine, 30% glycerol, 1 mM CHAPS.

Substrate (cleavage site)	K_m	k_{cat}	k_{cat}/K_m
	μM	s^{-1}	$M^{-1}s^{-1}$
Ac-DPNRKR-pNA (2A-2B)	740 ± 62	1.3 ± 0.04	1756 ± 96
Ac-LQYTKR-*p*NA (2B-3)	892 ± 95	1.1 ± 0.05	1233 ± 86
Ac-FASGKR-*p*NA (3–4A)	450 ± 45	1.9 ± 0.06	4222 ± 313
Ac-KPGLKR-*p*NA (4B-5)	821 ± 85	1.5 ± 0.06	1827 ± 124

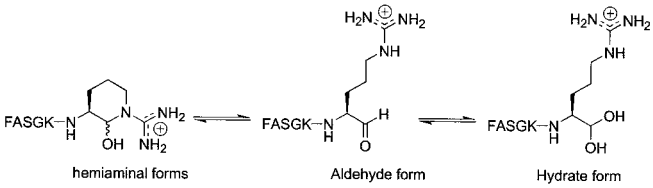

FIG. 6. **Chemical form of aldehyde inhibitor in water.**

form of the inhibitor (free aldehyde) available for interaction with the active site serine hydroxyl group.

The effect of standard protease inhibitors on processing of the substrate Ac-FASGKR-pNA by WNV CF40.gly.NS3pro was also investigated (Fig. 7). Standard serine protease inhibitors such as benzamidine, FUT-175, PMSF, leupeptin, and TLCK, failed to show any significant protease inhibition at concentrations of 100 μM. Under the same conditions FASGKR-aldehyde was a significant inhibitor even at a concentration of 10 μM (Fig. 7).

Homology Modeling—In order to provide a structural basis for further investigations of the WNV protease, for which no

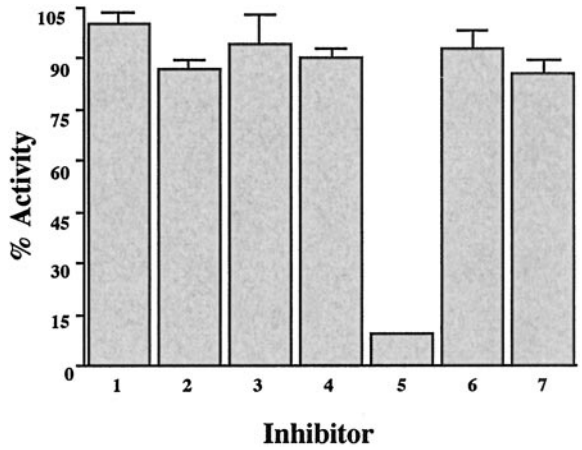

FIG. 7. **Inhibitor profile of WNV CF40.Gly.NS3pro.** A range of inhibitors were tested for their activity (±S.E.) against WNV CF40.Gly.NS3pro. Test compounds were preincubated with 0.5 μM WNV CF40.Gly.NS3pro for 10 min at 37 °C followed by addition of the substrate Ac-FASGKR-pNA substrate (500 μM) in 50 mM glycine-NaOH buffer, pH 9.0, 30% glycerol and 1 mM CHAPS. Inhibitors tested were: *1*, no inhibitor control; *2*, benzamidine (100 μM); *3*, FUT-175 (100 μM); *4*, PMSF (100 μM); *5*, FASGKR-aldehyde inhibitor (10 μM); *6*, leupeptin (100 μM); *7*, TLCK (100 μM).

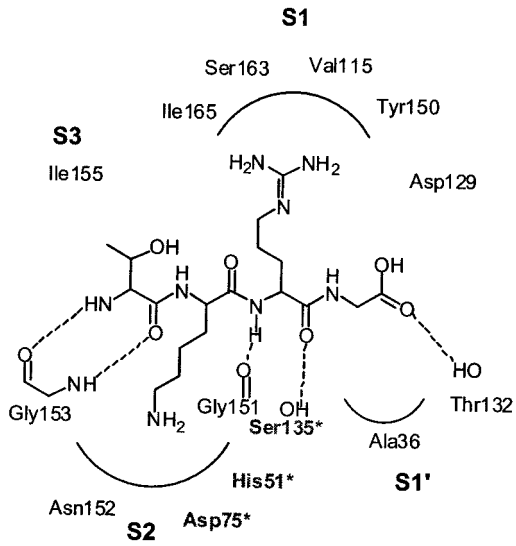

FIG. 8. **Homology-modeled structure of West Nile Virus NS3 protease showing putative interactions between WNV-NS3 protease active site residues (*S3*, *S2*, *S1*, *S1′*) and substrate.**

current crystal structure exists, a homology structural model was generated based on the crystal structures of two flavivirus proteases, the Dengue virus NS3 protease (pdb1BEF and pdb1DF9) (9), and the hepatitis C virus NS3 protease (pdb1A1Q) (8). The Dengue virus NS3 and hepatitis C virus proteases were selected as a basis for the WNV protease model for their sequence homology (with Dengue NS3pro, Fig. 2) and structural homology (HCV, with lower sequence homology), providing a suitable reference for generating a WNV NS3 protease homology model. As the flaviviral NS3 proteases require a cofactor for activity, the HCV protease structure (pdb1A1Q) (8) containing the HCV NS4A cofactor was used for the WNV protease homology model, thus taking into account the role of the WNV NS2B cofactor in the activation of the WNV NS3 protease.

Whereas sequence identity between the HCV protease and WNV protease or HCV protease and Dengue viral protease was much less (14.5% and 15.5% respectively), the knowledge of high structural homology based on a previous study of HCV and Dengue virus proteases allowed additional structural reference for the generation of the model. Studies by Brinkworth *et al.* (25) elucidated a model of Dengue virus NS3 protease based on the HCV NS3 protease crystal structure, despite low sequence identity. This has since been confirmed by a crystal structure for Dengue virus NS3 protease (pdb1DF9), used in the current WNV homology model construction. Sequence alignment of the HCV protease and WNV protease or HCV protease and Dengue virus protease yielded 50 and 59% similarity respectively, giving insight into the structural similarity between the three proteases.

The HCV model with NS4A cofactor allowed the generation of a WNV protease and potential cofactor model not seen in the Dengue virus crystal structure. The Dengue NS3 crystal structure in complex with mung bean Bowman-Birk inhibitor (pdb1DF9) illustrated the key interactions in the active site of the Dengue NS3 protease. The Dengue NS3 structure provided a useful reference for the development of WNV protease-substrate interaction studies. Knowledge of endogenous cleavage sites was later used in substrate docking and active site studies, with a view to designing an early substrate-based inhibitor of the West Nile Virus protease.

Numerous models were generated, based on comparisons between the different structures and using different sequence alignment techniques, in order to examine different possibilities for the WNV protease structure. The final model used for analysis of putative substrate interactions in the active site was based on the Dengue NS3 protease complexed with the Bowman-Birk inhibitor (pdb1DF9, subunit B). The α-carbon traces of the WNV model and the Dengue-2 viral protease crystal structure were very similar.

Based on the endogenous cleavage sites of the WNV NS3 protease, and their P1, P2 dibasic specificity, a four residue P3-P1′ substrate (Thr-Lys-Arg-Gly) was docked into the active site of the WNV protease model in order to analyze the probable binding interactions between the substrate and WNV protease (Fig. 8). When the Arg at P1 is docked into the protease S1 subsite, residues Tyr^{150}, Asp^{129}, Ser^{163}, Val^{115}, and Ile^{165} are in close proximity (Fig. 8). It seems likely that Arg makes an important pi-cation interaction with Tyr^{150} in the deep S1 pocket. The P2 substrate residue Lys can interact with protease S2 subsite residues Asn^{152}, Asp^{75}, Gly^{151}, and His^{51} (Fig. 8). The S2 subsite is fairly exposed, but a hydrogen bond may be formed between Lys at P2 and Asn^{152} at S2. The catalytic triad is also exposed and lies adjacent to the P2 pocket. The P3 Thr residue exhibited a potential interaction with protease subsite S3 residues Gly^{153} and Ile^{155} (Fig. 8). The P1′ Gly residue may interact with protease subsite S1′ by hydrogen bonding to Thr^{132} and is in close proximity to Ala^{36} (Fig. 8). Analysis of these interactions and their relative importance suggested possible mutagenesis of the active site residues Ser^{163}, Tyr^{150}, Asp^{129}, and Asn^{152}.[4]

DISCUSSION

West Nile Virus infections have increased dramatically over the past decade and the associated morbidity and mortality have increased the urgency for antiviral therapies (29). Rationally devised treatments for WNV infections might also become prototypes for developing antiviral therapies for other flavivirus infections. A single protease enzyme within the WNV genome is known to be responsible for critical virus-mediated post-translational cleavages of the polypeptide encoding all the structural and functional viral proteins that ultimately provide the replicative machinery for virus amplification and assembly

[4] K. J. Chappell, T. A. Nall, M. J. Stoermer, N.-X. Fang, J. D. A. Tyndall, D. P. Fairlie, and P. R. Young, manuscript in preparation.

into new virus particles. Inhibition of that protease has the capacity therefore to prevent viral replication. The present study was aimed at producing and characterizing the first catalytically active recombinant WNV protease.

To construct a recombinant West Nile Virus protease, we took advantage of the sequence homology between WNV and Dengue-2 virus NS2B-NS3 proteins (Fig. 2) and our previous work that successfully led to a recombinant Dengue virus protease (6). In summary, we adapted the corresponding WNV sequences of the putative minimum NS2B cofactor domain and of the minimum NS3 protease domain and linked them together with a nonapeptide (G_4SG_4) replacement for the NS2B-NS3A native cleavage site. The result was a catalytically active recombinant WNV protease, designated WNV CF40.Gly.-NS3pro, which is active only with the appended cofactor domain.

We characterized the enzyme processing properties of the recombinant WNV protease, identifying optimal buffers, pH, and diluents to maximize proteolytic activity. The assay, involving the use of hexapeptide substrates corresponding to P6-P1 segments of cleavage sites within the WNV genome fused to a chromogenic moiety (*para*-nitroanilide) at P1′, has the potential for rapidly screening prospective inhibitors of the enzyme. Indeed, based on one of the substrates examined, we also generated the first small molecule inhibitor (IC_{50} 1 μM) of the WNV protease.

A striking feature of the enzymatic activity was the similarity between catalytic properties of the recombinant proteases of Dengue virus and West Nile Virus. The high pH (8–10) required for processing is unusual and cannot yet be satisfactorily explained, but may indicate a level of virus-mediated control over replication, since processing at physiological pH would be 1% of the measured rates at pH 9.0.

The native cleavage sites that are processed by WNV protease *in vivo* (Fig. 1) feature two conserved basic residues at positions P1 and P2 (Arg, Lys, respectively, Fig. 4) and a small residue at P1′ (Gly, Ser), indicating a signal sequence for the cleavage of the WNV genome by its NS2B-NS3 protease. This dibasic recognition sequence is almost completely conserved throughout the flaviviral proteases and is unusual in comparison with host cellular enzymes. Indeed we have now shown for both the Dengue-2 (6) and WNV virus proteases (Fig. 7) that a panel of conventional serine protease inhibitors was inactive against that protease. This feature promises novelty for the design of potent inhibitors that may be selective for flavivirus proteases over host cellular proteases. Realization of a potent and selective inhibitor of WNV protease may result in an attractive lead for development of an antiviral treatment for WNV, and perhaps with potential for broad spectrum inhibition of flaviviruses in general.

Toward further studies, aimed at developing even more potent and selective inhibitors of WNV protease as candidate antiviral drugs, we also created a three-dimensional structural model of the protease. That model suggests binding sites in the protease for both the NS2B cofactor and small molecule substrates/inhibitors that can occupy the active site, and reveals putative enzyme-ligand interactions that should provide valuable insights for inhibitor development. This information could conceivably be combined with the known requirement in protease active sites of the extended β-strand ligand conformation (22, 27),[3] and the growing number of β-strand mimetics (30, 31) becoming available as templates for design of substrate-based inhibitors of proteases.

REFERENCES

1. Brinton, M. A. (2002) *Annu. Rev. Microbiol.* **56,** 371–402
2. Hayes, C. G. (2001) *Ann. N. Y. Acad. Sci.* **951,** 25–37
3. Lanciotti, 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., Cropp, 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. Teo, K. F., and Wright, P. J. (1997) *J. Gen. Virol.* **78,** 337–341
5. Buckley, A., Dawson, A., Moss, S. R., Hinsley, S. A., Bellamy, P. E., and Gould, E. A. (2003) *J. Gen. Virol.* **84,** 2807–2817
6. 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
7. Clum, S., Ebner, K. E., and Padmanabhan, R. (1997) *J. Biol. Chem.* **272,** 30715–30723
8. Love, R. A., Parge, H. E., Wickersham, J. A., Hostomsky, Z., Habuka, N., Moomaw, E. W., Adachi, T., Margosiak, S., Dagostino, E., and Hostomska, Z. (1998) *Clin. Diagn. Virol.* **10,** 151–156
9. Murthy, H. M., Clum, S., and Padmanabhan, R. (1999) *J. Biol. Chem.* **274,** 5573–5580
10. Gulick, R. M. (2003) *Clin. Microbiol. Infect.* **9,** 186–193
11. Rutenber, E., Fauman, E. B., Keenan, R. J., Fong, S., Furth, P. S., Ortiz de Montellano, P. R., Meng, E., Kuntz, I. D., DeCamp, D. L., Salto, R., Rose, J. R., Craik, C. S., and Stroud, R. M. (1993) *J. Biol. Chem.* **268,** 15343–15346
12. Lin, C., Lin, K., Luong, Y. P., Rao, B. G., Wei, Y. Y., Brennan, D. L., Fulghum, J. R., Hsiao, H. M., Ma, S., Maxwell, J. P., Cottrell, K. M., Perni, R. B., Gates, C. A., and Kwong, A. D. (2004) *J. Biol. Chem.* **279,** 17508–17514
13. Wuthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley-Interscience, New York
14. Amberg, S. M., and Rice, C. M. (1999) *J. Virol.* **73,** 8083–8094
15. Abbenante, G., Leung, D., Bond, T., and Fairlie, D. P. (2001) *LIPS* **7,** 347–351
16. Siev, D. V., and Semple, J. E. (2000) *Org. Lett.* **2,** 19–22
17. 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
18. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) *Nucleic Acids Res.* **22,** 4673–4680
19. Jones, D. T. (1999) *J. Mol. Biol.* **292,** 195–202
20. Hicks, R. P., Mones, E., Kim, H., Koser, B. W., Nichols, D. A., and Bhattacharjee, A. K. (2003) *Biopolymers* **68,** 459–470
21. Jones, G., Willett, P., and Glen, R. C. (1995) *J. Mol. Biol.* **245,** 43–53
22. 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., Burkett, B. A. (2000) *J. Med. Chem.* **43,** 1271–1281
23. Leung, D., Abbenante, G., Fairlie, D. P. (2000) *J. Med. Chem.* **43,** 305–341
24. Stocks, C. E., and Lobigs, M. (1998) *J. Virol.* **72,** 2141–2149
25. Brinkworth, R. I., Fairlie, D. P., Leung, D., and Young, P. R. (1999) *J. Gen. Virol.* **80,** 1167–1177
26. Arakaki, T. L., Fang, N. X., Fairlie, D. P., Young, P. R., and Martin, J. L. (2002) *Protein Expr. Purif.* **25,** 241–247
27. Tyndall, J. D. A., and Fairlie, D. P. (1999) *J. Mol. Recognit.* **12,** 363–370
28. Schultz, R. M., Varma-Nelsons, P., Ortiz, R., Kozlowskin, K. A., Orawski, A. T., Pagast, P., and Frankfater, A. (1989) *J. Biol. Chem.* **264,** 1497-
29. Campbell, G. L., Marfin, A. A., Lanciotti, R. S., and Gubler, D. J. (2002) *Lancet Infect. Dis.* **2,** 519–529
30. Glenn, M. P., and Fairlie, D. P. (2002) *Mini Rev. Med. Chem.* **2,** 433–445
31. Loughlin, W. A., Tyndall, J. D. A., Kelso, M. K., and Fairlie, D. P. (2004) *Chem. Reviews*, in press
32. Falgout, B., Miller. R. H., and Lai, C. J. (1993) *J. Virol.* **67,** 2034–2042

Enzymatic Characterization and Homology Model of a Catalytically Active Recombinant West Nile Virus NS3 Protease

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

J. Biol. Chem. 2004, 279:48535-48542.
doi: 10.1074/jbc.M406810200 originally published online August 18, 2004
