Protein Science (1998), *7*:2143–2149. Cambridge University Press. Printed in the USA.
Copyright © 1998 The Protein Society

Construction, expression, and characterization of a novel fully activated recombinant single-chain hepatitis C virus protease

S. SHANE TAREMI, BRIAN BEYER, MAUREEN MAHER, NANHUA YAO, WINIFRED PROSISE, PATRICIA C. WEBER, AND BRUCE A. MALCOLM
Department of Structural Chemistry, Schering-Plough Research Institute, Kenilworth, New Jersey 07033

(RECEIVED August 13, 1998; ACCEPTED August 17, 1998)

Abstract

Efficient proteolytic processing of essential junctions of the hepatitis C virus (HCV) polyprotein requires a heterodimeric complex of the NS3 bifunctional protease/helicase and the NS4A accessory protein. A single-chain recombinant form of the protease has been constructed in which NS4A residues 21–32 (GSVVIVGRIILS) were fused in frame to the amino terminus of the NS3 protease domain (residues 3–181) through a tetrapeptide linker. The single-chain recombinant protease has been overexpressed as a soluble protein in *E. coli* and purified to homogeneity by a combination of metal chelate and size-exclusion chromatography. The single-chain recombinant protease domain shows full proteolytic activity cleaving the NS5A-5B synthetic peptide substrate, DTEDVVCCSMSYTWTGK with a K_m and k_{cat} of 20.0 ± 2.0 μM and 9.6 ± 2.0 min^{-1}, respectively; parameters identical to those of the authentic $NS3_{1-631}/NS4A_{1-54}$ protein complex generated in eukaryotic cells (Sali DL et al., 1998, *Biochemistry 37*:3392–3401).

Keywords: hepatitis C virus; kinetics; NS3; NS4A; overexpression; protease; recombinant

Hepatitis C virus (HCV) is the major etiologic agent of non-A non-B (NANB) hepatitis with an estimated human seroprevalence of 1% globally (Alter & Mast, 1994). The majority of infections result in chronic hepatitis with a significant fraction of those progressing to cirrhosis and hepatocellular carcinoma (Stadhouders & Cooreman, 1997). The current therapeutic regimens are effective in less than 50% of the cases and consequently new anti-HCV agents are actively being sought (Sharara, 1997).

HCV, like all members of the *flaviviridae*, is a (+)-stranded RNA virus whose genome encodes a single polyprotein of about 3,000 amino acids which undergoes subsequent proteolysis by host and viral enzymes to yield the mature viral proteins (for reviews see Lohmann et al., 1996; Simmonds, 1996; Neddermann et al., 1997; Reed & Rice, 1998). The virally encoded, highly specific NS3 proteolytic activity is responsible for the essential processing of the NS3-4A, NS4A-NS4B, NS4B-NS5A, and NS5A-NS5B junctions (Grakoui et al., 1993; Tomei et al., 1993; D'Souza et al., 1994; Failla et al., 1994; Lin et al., 1994) and consequently is a primary target for antiviral therapeutics. The proteolytic domain, a twin-β-barrel serine protease of the chymotrypsin family, constitutes the amino terminal third of the NS3 protein (residues 1–181) (Hahm et al., 1995; Kim et al., 1996; Love et al., 1996; Yan et al., 1998). A key feature of the NS3 protease is that one strand of its seven stranded N-terminal β-barrel structure is supplied in trans by the NS4A cofactor protein or peptide. Upon binding of the cofactor peptide, the disordered N-terminus of NS3 (residues 1–28), folds to generate a new β strand (A0) and an α helix (α0) that encase residues 21–33 of NS4A (Tanji et al., 1995; Kim et al., 1996; Yan et al., 1998). This causes the NS4A strand to reposition the D1 strand of the N-terminal β barrel, which in turn stabilizes the α1 helix that bears the second member of the catalytic triad, His57 (Tanji et al., 1995; Kim et al., 1996; Yan et al., 1998). In addition, the refolding of the N-terminus causes the side chain of Asp81 to reorient from the guanidinium group of Arg155 to the imidazole ring of His57 to generate a properly arranged catalytic triad (Tanji et al., 1995; Kim et al., 1996; Yan et al., 1998). In summary, without intercalation of the NS4A strand the N-terminus remains partially disordered resulting in imperfect alignment of the catalytic triad and a corresponding ~950-fold drop in catalytic efficiency (Tanji et al., 1995; Kim et al., 1996; Yan et al., 1998).

Reprint requests to: Bruce A. Malcolm, Schering-Plough Research Institute, 2015 Galloping Hill Road, Kenilworth, New Jersey 07033; e-mail: bruce.malcolm@spcorp.com.

Abbreviations: aa, amino acid; βOG, n-octyl-β-D-octyl glucoside; βME, β-mercaptoethanol; BSA, bovine serum albumin; DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid; ESI-MS, electrospray ionization mass spectrometry; HCV, hepatitis C virus; HEPES, *N*-(2-hydroxyethyl)piperizine-*N'*-2-ethanesulfonic acid; HPLC, high-performance liquid chromatography; IPTG, isopropyl thio-β-D-galactoside; LB, Luria broth; MES, 2-(*N*-morpholino)ethanesulfonic acid; MOPS, 3-(*N*-morpholino)propanesulfonic acid; NMR, nuclear magnetic resonance spectroscopy; NS, nonstructural; nt, nucleotide; PAGE, polyacrylamide gel electrophoresis; PCR, polymerase chain reaction; SDS, sodium dodecyl sulfate; TFA, trifluoroacetic acid.

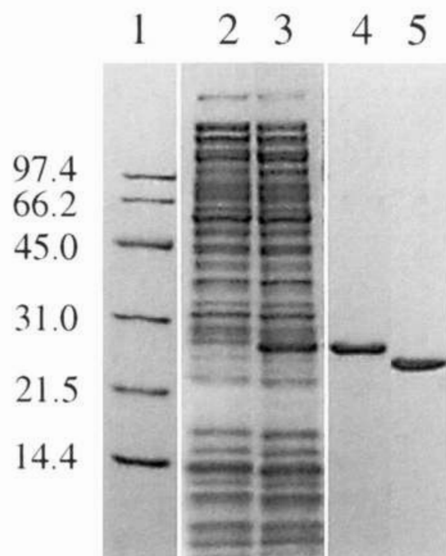

Fig. 1. SDS-PAGE analysis (Coomassie stained) of $NS4A_{21-32}$–GSGS–$NS3_{3-181}$ protein purification. Lane 1, molecular weight standards; lane 2, uninduced cleared cell lysate; lane 3, induced cleared cell lysate; lane 4, pooled fractions following Ni-imidoacetate purification (with histidine tag leader sequence); lane 5, pooled fractions following thrombin digestion and Sephacryl-100 chromatography.

Various strategies have been attempted over the last few years to generate milligram quantities of this key target enzyme in a fully active, tractable form for assay and structural analysis. Although modestly successful, expression of the full $NS3_{1-631}/NS4A_{1-54}$ heterodimeric protein complex in eukaryotic systems yielded relatively small quantities of a rather insoluble material insufficient in quantity or quality for structural analysis (Hirowatari et al., 1993; Overton et al., 1995; Suzuki et al., 1995; Steinkuhler et al., 1996; Morgenstern et al., 1997; Sali et al., 1998). Examination of the NS4A sequence suggested that the amino-terminal portion (residues 1–21) was primarily responsible for the lack of solubility of the $NS3_{1-631}/NS4A_{1-54}$ protein complex and subsequently synthetic peptides, based only on the intercalating portion of NS4A (residues 22–33), were used in place of the NS4A protein (Butkiewicz et al., 1996; Tomei et al., 1996). Proteolytic activity, however, remained an order of magnitude lower than the $NS3_{1-631}/NS4A_{1-54}$ protein complex despite large molar excesses of these peptides (Sali et al., 1998).

Expression of the NS3 protease domain by itself (residues 1–181) in *E. coli* yielded milligram quantities of material suitable for structural analysis; however, again reduced proteolytic activity was observed even in the presence of excess cofactor peptide (Shoji et al., 1995; Vishnuvardhan et al., 1997). Biophysical studies have suggested that the relatively low micromolar affinity of these peptides for the protease were most likely responsible for the poor trans complementation and partial restoration of the proteolytic activity (Bianchi et al., 1997).

To generate a reagent suitable for both structural analysis and high throughput screening that circumvents the addition of an exogenous *trans* acting peptide cofactor, a single-chain construct encoding the requisite peptide from the NS4A cofactor has been fused in frame to the NS3 protease domain. This construct, presented below, can be overexpressed in *E. coli* to high levels as a soluble protein and retains full catalytic activity compared with the $NS3_{1-631}/NS4A_{1-54}$ protein complex produced in eukaryotic cells.

Results

Recombinant single-chain $NS4A_{21-32}$–GSGS–$NS3_{3-181}$ constituted approximately 5% of the intracellular protein following induction as shown in Figure 1. Following purification, yields approached 10 mg/L of culture. ESI-MS gave a monoisotopic mass of 20,753.0 ± 2.5 Da, which is in good agreement with the calculated value of 20,751.4 Da. Edman sequencing revealed the anticipated sequence beginning with the P′ portion of the thrombin cleavage sequence (GSHM) and proceeding to that of the NS4A peptide (GSVVIVGRII).

Kinetic analysis

Kinetic parameters for the recombinant single-chain $NS4A_{21-32}$–GSGS–$NS3_{3-181}$ protease were obtained from nonlinear regression analysis of Michaelis–Menten plots (Fig. 2). A comparison of k_{cat} and K_m of the recombinant single-chain $NS4A_{21-32}$–GSGS–$NS3_{3-181}$, with that of the full length $NS3_{1-631}/NS4A_{1-54}$ protein complex produced in the baculovirus system (Sali et al., 1998) and the recombinant NS3 catalytic domain complemented in *trans* with an excess of the synthetic peptide GSVVIVGRIILS ($NS3_{1-181}$ + NS4A peptide), is shown in Table 1. The kinetic parameters for the

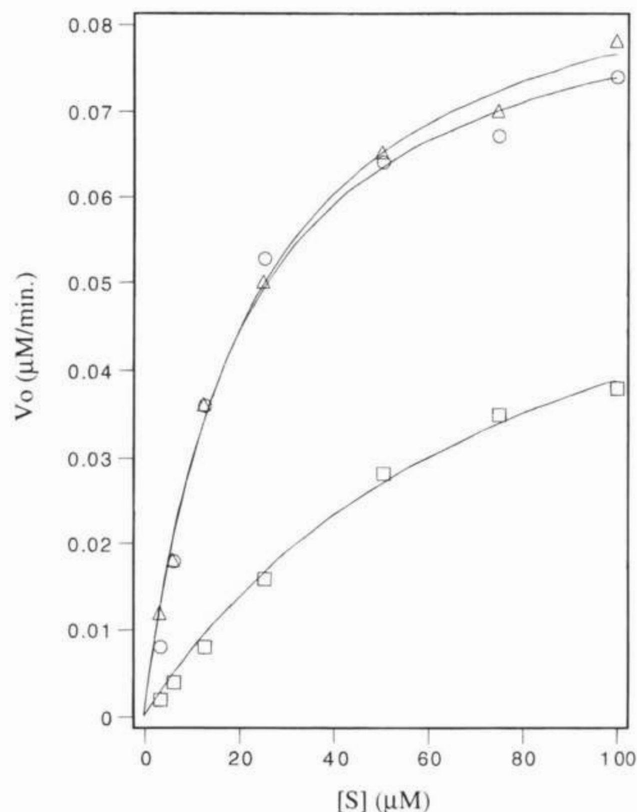

Fig. 2. Dependence of initial velocity on substrate concentration for the recombinant constructs described in Table I (□ = $NS3_{1-181}$ + NS4A peptide, ○ = $NS4A_{21-32}$–GSGS–$NS3_{3-181}$, △ = $NS3_{1-631}NS4A_{1-54}$). Graph shows exemplary data sets for each construct (see Materials and methods).

Table 1. *Kinetic parameters of HCV protease constructs*[a]

Construct	k_{cat} (min^{-1})	K_m (μM)	k_{cat}/K_m ($M^{-1} s^{-1}$)
$NS3_{1-631}/NS4A_{1-54}$	10 ± 2	20 ± 2	(8 ± 2) × 10^3
$NS3_{1-181}$ + NS4A peptide	3 ± 1	80 ± 20	(0.5 ± 0.2) × 10^3
$NS4A_{21-32}$–GSGS–$NS3_{3-181}$	9 ± 2	19 ± 3	(8 ± 2) × 10^3

[a]5A–5B cleavage junction substrate DTEDVVCC*SMSYTWTG-K (see Materials and methods).

catalytic domain generated in *E. coli* and complemented with a 1,000-fold excess of the NS4A peptide are measurably different (k_{cat} = 3 min^{-1}; K_m = 80 μM) from those of the single-chain recombinant $NS4A_{21-32}$–GSGS–$NS3_{3-181}$ catalytic domain and the full length $NS3_{1-631}/NS4A_{1-54}$ protein complex generated in the insect cell system (k_{cat} ~ 9–10 min^{-1}; K_m ~ 19–20 μM). These differences suggest a greater than 15-fold improvement in the second order rate constant as a consequence of tethering the cofactor to the catalytic domain.

The pH profile (Fig. 3A) and the βOG profile (Fig. 3D) of the single-chain construct ($NS4A_{21-32}$–GSGS–$NS3_{3-181}$) suggests a pH optimum of 7.5 and a detergent optimum of 0.05%, which are comparable to that of the full $NS3_{1-631}/NS4A_{1-54}$ protein complex from the baculovirus expression system (studies reported in Sali et al., 1998). Unlike the full length complex however (see data in Sali et al., 1998), the single-chain construct shows decidedly higher ionic strength and glycerol optima (Fig. 3B,C). Over the range of the studies (0–2 M NaCl and 0–50%, respectively), there appears to be a simple monotonic increase in activity with increasing ionic strength and concentration of glycerol.

Discussion

Both kinetic and structural studies have shown that the HCV NS3 serine protease requires intercalation of a strand of the NS4A cofactor for proper alignment of the catalytic triad and full proteolytic activity (Kim et al., 1996; Yan et al., 1998). Two approaches have been used to obtain the fully active target enzyme: (1) production of a full length $NS3_{1-631}/NS4A_{1-54}$ complex in eukaryotic expression systems (Hirowatari et al., 1993; Overton et al., 1995; Suzuki et al., 1995; Steinkuhler et al., 1996; Morgenstern et al., 1997; Sali et al., 1998) and (2) production of the protease domain alone in *E. coli* followed by complementation in *trans* with an excess of synthetic oligopeptide (Shoji et al., 1995; Vishnuvardhan et al., 1997).

The former approach, although feasible, is labor intensive, expensive, and ultimately provides a complex ill-suited to structural analysis by either X-ray diffraction or NMR spectroscopy due to

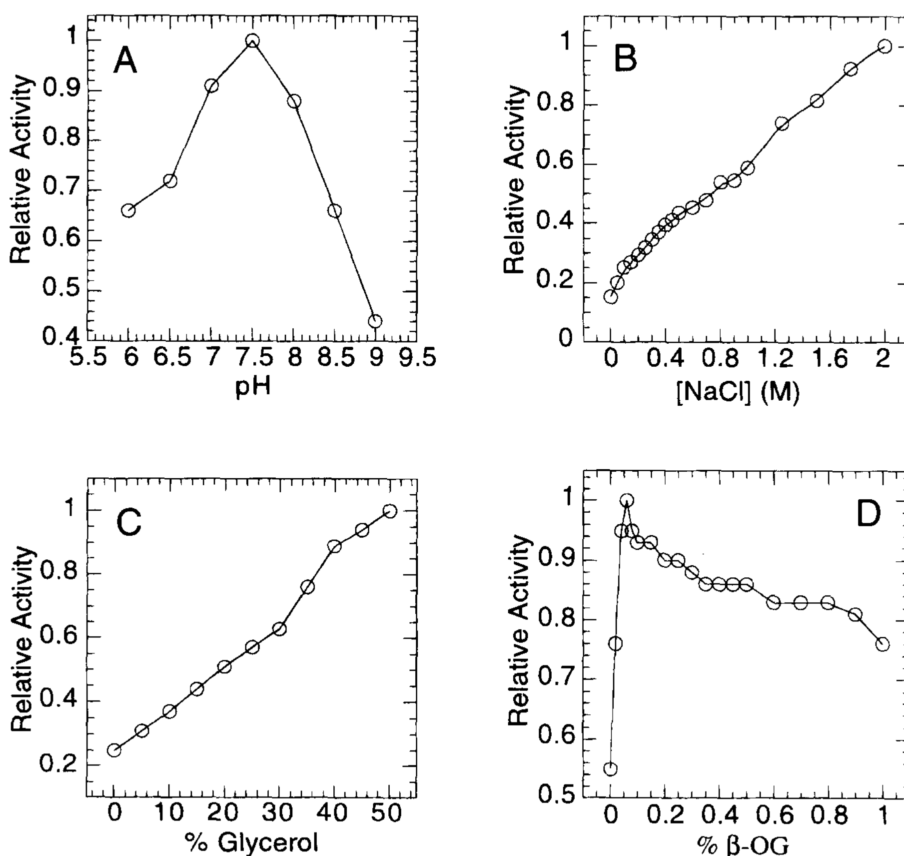

Fig. 3. Dependence of $NS4A_{21-32}$–GSGS–$NS3_{3-181}$ activity on (**A**) pH, (**B**) ionic strength, (**C**) concentration of glycerol, and (**D**) concentration of detergent-βOG. For experimental details, see Materials and methods.

size and solubility. The latter method, although reasonably fast, inexpensive, and capable of generating material suitable for structural analysis, has been criticized for the low specific activity of the resulting material compared with the full length $NS3_{1-631}$/$NS4A_{1-54}$ protein complex (Kim et al., 1996; Sali et al., 1998; Yan et al., 1998). Recent studies have shown that the NS4A core peptide's affinity for the protease is several orders of magnitude less than that of the full NS4A protein (Kim et al., 1996; Bianchi et al., 1997; Sali et al., 1998; Yan et al., 1998). Rather than attempt to modify the NS4A peptide cofactor to improve its intrinsic affinity for NS3, a fusion strategy to anchor the cofactor to the protease was devised.

Inspection of the cocrystal structure of the $NS3_{1-181}$/$NS4A_{21-39}$ complex suggested that it might be possible to fuse the C-terminus of the essential portion of the NS4A peptide (residues 21–32) to the N-terminus of the catalytic domain using an appropriate spacer (Fig. 4). It was reasoned that if an appropriate connection could be engineered, then the tethered cofactor peptide could not undergo complete dissociation and the single-chain construct would more likely remain in the completely ordered state (Kim et al., 1996; Yan et al., 1998). The result would be a small soluble NS3 protease with full catalytic activity requiring no exogenous cofactor. Various linkers, some intentionally unstructured and others incorporating actual turn sequences from known structures (e.g., PKG, PAP, PAG, etc.), were considered to connect the nearest residues of the C-terminal region of the NS4A cofactor and the N-terminal region of the NS3 protease that appeared to retain structural importance—Ser32 and Ile3, respectively (Kim et al., 1996; Yan et al., 1998).

After evaluation by visual inspection and energy minimization, four linkers—PAG, PAGG, GSG, and GSGS—were chosen for incorporation into constructs. Although the PAG, PAGG, and GSG containing constructs showed some degree of soluble expression ($\leq$1% of the intracellular protein) and comparable specific activity (data not shown), the flexible GSGS linker was found to yield the highest levels (~5% of the intracellular protein) and consequently was chosen for further study. Kinetic characterization suggests that the catalytic efficiency of the single-chain catalytic domain construct is greater than the NS3 catalytic domain complemented in *trans* by the equivalent NS4A peptide ($NS3_{1-181}$ + NS4A peptide) and identical to that of the full length $NS3_{1-631}$/$NS4A_{1-54}$ protein complex. Although identical interactions would be expected between the backbone and side chains of the intercalated portion of the cofactor and the protease domain in all cases, it is reasonable to assume that covalently restraining the peregrinations of the C-terminus of the peptide (not to mention preventing complete dissociation) must reduce disorder in the system with respect to the $NS3_{1-181}$ + NS4A peptide system (whose K_d has been estimated to be 3.8–7.7 μM (Bianchi et al., 1997)), yielding a more stable active site comparable to that of the full length complex (for which the dissociation appears to be sub-nanomolar (Sali et al., 1998)).

The similarities and differences of the single-chain proteolytic domain and the full length $NS3_{1-631}$/$NS4A_{1-54}$ protein complex with respect to global parameters such as pH, ionic strength, glycerol, and detergent optima can be understood through their common and distinct elements. The fact that both proteases exhibit the same pH optimum is not surprising, as this parameter directly reflects the ionization and structural alignment of their catalytic machineries (Perona & Craik, 1995 and references therein), which appear, from kinetic analyses, to be identical. The comparable effect of detergent and low levels of salt and glycerol on both systems likewise suggests that this is the result of impact on the common elements; either the proteolytic domain (including the intercalated portion of NS4A) and/or peptide substrate binding. The deleterious result of high ionic strength and glycerol on only the NS3/NS4A protein complex suggests that these negative effects are mediated by the distinct elements peculiar to the NS3/NS4A protein complex; specifically the nonintercalated portions of NS4 (residues 1–21 and 33–54) and the helicase domain of NS3 (residues 182–631). The single-chain construct, lacking these pieces, is therefore refractory.

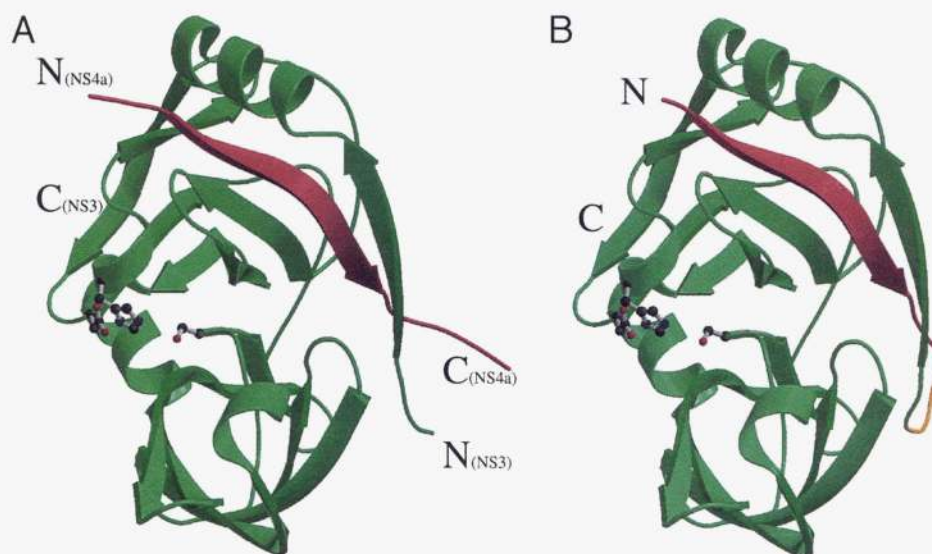

Fig. 4. Ribbon diagram of the proteolytic domain of NS3 showing the tethering of the NS4A cofactor. **A:** X-ray structure of $NS3_{1-181}$ + $NS4A_{20-35}$ peptide. **B:** A model of $NS4A_{21-32}$–GSGS–$NS3_{3-181}$. The intercalated portion of the construct corresponding to the NS4A cofactor peptide (residues 21–32) is shown in red and the GSGS linker in yellow. The catalytic triad (Ser139, His57, and Asp81) side chains are shown as ball and stick models.

Although there may be subtle differences between the full $NS3_{1-631}/NS4A_{1-54}$ protein complex and the single-chain proteolytic domain ($NS4A_{21-32}$–GSGS–$NS3_{3-181}$) with regard to processing of polyprotein substrates, prima facie kinetic evidence suggests that the single-chain construct contains an optimally organized active site identical to that of the $NS3_{1-631}/NS4A_{1-54}$ protein complex and should show a similar specificity profile with these substrates. Structural and polyprotein cleavage studies are in progress to confirm these predictions. In the interim, this novel single-chain construct should allow the facile generation of reagent quantities of this important biomedical target and facilitate the identification of medicinally useful inhibitors.

Materials and methods

Molecular biology

Restriction enzymes, Vent polymerase, and ThermoPol buffer were obtained from New England Biolabs (Beverly, Massachusetts). PCR reactions were carried out in a Perkin Elmer (Foster City, California) Cetus, model 480 DNA thermocycler. pET vectors and BL21(DE3) cells were obtained from Novagen (Madison, Wisconsin). DH5α cells were purchased from Gibco, BRL (Gaithersburg, Maryland). Clone BK 138-1, described in Takamizawa et al. (1991) encoding the entire NS3 region of the 1b/BK strain of HCV, was obtained from the authors. Standard recombinant DNA methods were carried out essentially as described by Sambrook et al. (1989). DNA sequencing was performed according to the Sanger-Dideoxy method by Bioserve Biotechnologies (Laurel, Maryland).

Protein analysis

Protein concentrations were determined by the BIORAD dye method against recombinant HCV standards previously quantified using amino acid analysis. N-terminal sequence analysis was performed on an Applied Biosystems Inc. (Foster City, California), 494 pulsed-liquid protein sequencer with an on-line PTH analyzer (ABI model 140C). Samples were spotted onto PVDF membranes (ProSorb, ABI, CA) and were subjected to a minimum of 10 cycles of analysis. Electrospray ionization mass spectrometry (ESI-MS) was performed on a Sciex API III triple quadrupole mass spectrometer (Ontario, California) fitted with a pneumatically assisted electrospray (ionspray) device. The sample solution (25–50 μM in 5 mM ammonium acetate, pH 6.5) was infused at 5 μL/min with an orifice voltage of 80–100 V. Multiple scans from 1,650–2,400 m/z were acquired for each sample. The standard deviation of the mass estimates ranged from 2–3 Da.

Peptide synthesis

The P8P'8K substrate peptide derived from the 5A–5B junction of the 1a/H strain of HCV was synthesized using standard Fmoc solid phase synthesis protocols on an Applied Biosystems Inc., 431 A Peptide Synthesizer as previously described (Zhang et al., 1997). The substrate was purified using reverse-phase HPLC and the mass confirmed by plasma desorption mass spectrometry using a Bioion 20 Californium-252 time-of-flight instrument as before.

Assay of protease activity

The activity of protease constructs was determined using the peptide cleavage assay previously described (Sali et al., 1998). In brief, varying amounts of the 5A/5B (P8P'8) substrate peptide DTEDVVCC*SMSYTWTG-K (the * denotes the scissile bond; the C-terminal K was added to facilitate separation of the product) were incubated with appropriate amounts of the recombinant proteases in 25 mM Tris-HCl (pH 7.5), 150 mM NaCl, 0.5 mM EDTA, 10 mM DTT, 10% glycerol, and 0.05% βOG at room temperature for 15 min and subsequently quenched with an equal volume of 10% phosphoric acid. Activity profiles varying ionic strength, percent glycerol, and concentration of βOG were conducted in the above buffer. The pH profile study was conducted using the standard solution conditions above with the following buffers: 25 mM MES-HCl, pH 6.0–6.5, 25 mM MOPS-HCl, pH 7.0, and 25 mM Tris-HCl, pH 7.5–9.0.

Cleavage products were quantified by reverse-phase HPLC. Peak areas for the P' product (P'1P'8-K) were converted to nanomoles using a standard curve calibrated using authentic material. Initial rates were determined from samples showing $\leq$10% substrate cleavage. Kinetic parameters were obtained using nonlinear regression analysis of initial rates as a function of substrate concentration (Enzyme Kinetics; Trinity Software, Campton, New Hampshire). All parameters represent the weighted average of three independent determinations.

Recombinant NS3 catalytic domain ($NS3_{1-181}$) without the fused NS4A peptide was preincubated with synthetic NS4A peptide (KKGSVVIVGRIVLSGKPAIIPKK) for 15 min at 4 °C to generate the activated complex ($NS3_{1-181}$ + NS4A peptide) prior to kinetic analysis (NS3 final concentration 250 nM, NS4A peptide 20 μM). The mixture was then diluted directly into the proteolytic reaction to give a final concentration of 100 nM NS3 catalytic domain/8 μM NS4A peptide. Full length $NS3_{1-631}/NS4A_{1-54}$ protease was prepared as previously described (Sali et al., 1998).

Construction of recombinant single-chain $NS4A_{21-32}$–GSGS–$NS3_{3-181}$

Initial DNA sequencing of the BK 138-1 plasmid harboring the entire NS3 bifunctional protein revealed four nonsynonymous differences with the published sequence (Takamizawa et al., 1991), at positions aa 66/nt 3606 (A $\rightarrow$ G; gcg $\rightarrow$ ggc), aa 86/nt 3666 (P $\rightarrow$ Q; ccc $\rightarrow$ cag), aa 87/nt 3669 (K $\rightarrow$ A; aag $\rightarrow$ gcg), and aa 147/ nt 3879 (F $\rightarrow$ S; ttc $\rightarrow$ tcc) of the protease catalytic domain. A 5' PCR primer encoding: (1) an Nde I site, (2) residues 21–32 of NS4A, and (3) a tetrapeptide linker, Gly-Ser-Gly-Ser along with residues 3–9 of NS3, in frame with the pET 28b+ cistron, 5'GATATACATATGGGTTCTGTTGTTATTGTTGGTA GAATTATTTTATCTGGTAGTGGTAGTATCACGGCCTACTCC CAA-3' was used along with a 3' PCR primer complementary to amino acid residues 175–181 of NS3 as well as encoding an in frame opal termination codon flanked by an EcoRI site 5'-CTCAG CGAATTCTCAAGACCGCATAGTAGTTTCCAT-3' to amplify (in frame) residues 3–181 of plasmid BK 138-1. To improve fidelity, Vent DNA polymerase was used following the manufacturer's recommendations to amplify the DNA. Final primer/template/dNTP concentrations were: primers 0.1 μM; template 0.1 μg/mL; dNTPs 2.5 mM each. The PCR conditions were as follows: 95 °C for 45 s (pre-cycle); 95 °C for 30 s, 55 °C for 1 min, 72 °C for 2 min (25 cycles).

The amplified 616 base pair fragment was purified and double digested with EcoRI and Nde I, as were aliquots of the expression vector, pET-28b+. The digested fragment and vector were then gel purified, ligated overnight at 15 °C, and used to transform compe-

tent DH5α cells according to the manufacturer's protocol. Transformed cells were selected on LB agar plates with kanamycin (25 μg/mL). Recombinant clones were identified by restriction analysis and three were selected for complete sequence verification. The resulting recombinant single-chain constructs fused the NS4A cofactor (residues 21–32), the GSGS linker, and the NS3 protease (residues 3–181) followed by an opal stop codon, in frame to the histidine leader sequence (MGSSHHHHHHSSGLVPRG SHM) of pET-28b+. An identical construct, but lacking the NS4A cofactor and linker, was assembled to generate the catalytic domain alone ($NS3_{1-181}$).

Expression and purification

Both the single-chain proteolytic domain ($NS4A_{21-32}$–GSGS–$NS3_{3-181}$) and the catalytic domain alone ($NS3_{1-181}$) were expressed and purified following the procedure described below. Single colonies from plates of freshly transformed BL21(DE3) cells were used to initiate growth of all recombinant protease constructs. Terrific Broth medium was supplemented with 25 μg/mL kanamycin. When the cell density reached an OD_{600} of 2–3, the fermentation temperature was lowered rapidly to 23 °C and recombinant protein expression induced with IPTG (0.4 mM final concentration). Cells were harvested 3 h after induction and frozen at −20 °C prior to purification.

Cell pellets were resuspended in 600 mL of lysis buffer containing 50 mM HEPES, pH 7.4, 10% glycerol, 0.3 M NaCl, 0.1% βOG, 2 mM βME (buffer A), using a cell homogenizer (Omni Mixer ES) for 2 min and subsequently disrupted by two passes through a Microfluidizer, Model #M-110F (Microfluidics, Newton, Massachusetts) at 10,000 psi. The lysate was cleared by centrifugation at 85,000 × *g* for 45 min. The supernatant was then filtered through 0.8 micron filter units (Nalgene) and applied at 40 mL/min to a 11 mL Ni-imidodiacetate (POROS 20 MC resin) column in the presence of 20 mM imidazole using a BIOCAD system (Perseptive Biosystems, Framingham, Massachusetts). The column was washed with 10 column volumes of buffer A, followed by 15 column volume of buffer A containing 1.0 M NaCl and 20 mM imidazole (buffer B). The bound protease was eluted using buffer A supplemented with 1 M NaCl and 250 mM imidazole.

The eluted fractions containing the protease were pooled and human thrombin (Enzyme Research, South Bend, Indiana) was added to yield a thrombin to recombinant protease ratio of 8 U/mg of HCV enzyme. Digestion was allowed to proceed throughout the subsequent dialysis step. The thrombin/recombinant HCV protease mixture was dialyzed at 4 °C for 18 h against 16 L of 50 mM HEPES, pH 7.4, 10% glycerol, 1 M NaCl, and 10 mM βME to remove the imidazole and the detergent.

Thrombin was subsequently removed by affinity chromatography using 1 mL heparin sepharose column (Pharmacia, Piscataway, New Jersey). The pooled fractions were adjusted to 0.3 M NaCl, 25 mM HEPES (pH 7.4), 10% glycerol, and 10 mM βME prior to application to the column, previously equilibrated in the same buffer. The flow through was concentrated to 5 mg/mL using Centriprep-30 concentrators (Amicon, Beverly, Massachusetts) and further purified by size exclusion chromatography using 3 Sephacryl-100 sizing columns (26 × 60 cm, Pharmacia) in series (flow rate 0.5 mL/min). All columns were pre-equilibrated in buffer containing 50 mM HEPES, pH 7.4, 10% glycerol, 1 M NaCl, and 10 mM βME. Fractions containing greater than 95% pure recombinant protease, as judged by SDS-PAGE, were pooled and flash-frozen at −80 °C.

Acknowledgments

The authors would like to thank Drs. Yan-hui Liu, Charles McNemar, and Rumin Zhang for their support with mass spectrometry, protein sequencing, and peptide synthesis, respectively, and acknowledge the expert technical assistance of Hernando de Fex, James Durkin, and Richard Ingram in the preparation of peptide substrates and protein complexes.

References

Alter MJ, Mast EE. 1994. The epidemiology of viral hepatitis in the United States. *Gastroenterol Clin North Am 23*:437–455.

Bianchi E, Urbani A, Biasiol G, Brunetti M, Pessi A, De Francesco R, Steinkuhler C. 1997. Complex formation between the hepatitis C virus serine protease and a synthetic NS4A cofactor peptide. *Biochemistry 36*:7890–7897.

Butkiewicz NJ, Wendel M, Zhang R, Jubin R, Pichardo J, Smith EB, Hart AM, Ingram R, Durkin J, Mui PW, Murray MG, Ramanathan L, Dasmahapatra B. 1996. Enhancement of hepatitis C virus NS3 proteinase activity by association with NS4A-specific synthetic peptides: Identification of sequence and critical residues of NS4A for the cofactor activity. *Virology 225*:328–338.

D'Souza ED, O'Sullivan E, Amphlett EM, Rowlands DJ, Sangar DV, Clarke BE. 1994. Analysis of NS3-mediated processing of the hepatitis C virus non-structural region in vitro. *J Gen Virol 75*:3469–3476.

Failla C, Tomei L, De Francesco R. 1994. Both NS3 and NS4A are required for proteolytic processing of hepatitis C virus nonstructural proteins. *J Virol 68*:3753–3760.

Grakoui A, McCourt DW, Wychowski C, Feinstone SM, Rice CM. 1993. Characterization of the hepatitis C virus-encoded serine proteinase: Determination of proteinase-dependent polyprotein cleavage sites. *J Virol 67*:2832–2843.

Hahm B, Han DS, Back SH, Song OK, Cho MJ, Kim CJ, Shimotohno K, Jang SK. 1995. NS3-4A of hepatitis C virus is a chymotrypsin-like protease. *J Virol 69*:2534–2539.

Hirowatari Y, Hijikata M, Tanji Y, Nyunoya H, Mizushima H, Kimura K, Tanaka T, Kato N, Shimotohno K. 1993. Two proteinase activities in HCV polypeptide expressed in insect cells using baculovirus vector. *Arch Virol 133*:349–356.

Kim JL, Morgenstern KA, Lin C, Fox T, Dwyer MD, Landro JA, Chambers SP, Markland W, Lepre CA, O'Malley ET, Harbeson SL, Rice CM, Murcko MA, Caron PR, Thomson JA. 1996. Crystal structure of the hepatitis C virus NS3 protease domain complexed with a synthetic NS4A cofactor peptide [published erratum appears in *Cell 89*(1):159]. *Cell 87*:343–355.

Lin C, Pragai BM, Grakoui A, Xu J, Rice CM. 1994. Hepatitis C virus NS3 serine proteinase: Trans-cleavage requirements and processing kinetics. *J Virol 68*:8147–8157.

Lohmann V, Koch JO, Bartenschlager R. 1996. Processing pathways of the hepatitis C virus proteins. *J Hepatol 24*:11–19.

Love RA, Parge HE, Wickersham JA, Hostomsky Z, Habuka N, Moomaw EW, Adachi T, Hostomska Z. 1996. The crystal structure of hepatitis C virus NS3 proteinase reveals a trypsin-like fold and a structural zinc binding site. *Cell 87*:331–342.

Morgenstern KA, Landro JA, Hsiao K, Lin C, Gu Y, Su MS, Thomson JA. 1997. Polynucleotide modulation of the protease, nucleoside triphosphatase, and helicase activities of a hepatitis C virus NS3-NS4A complex isolated from transfected COS cells. *J Virol 71*:3767–3775.

Neddermann P, Tomei L, Steinkuhler C, Gallinari P, Tramontano A, De Francesco R. 1997. The nonstructural proteins of the hepatitis C virus: Structure and functions. *Biol Chem 378*:469–476.

Overton H, McMillan D, Gillespie F, Mills J. 1995. Recombinant baculovirus-expressed NS3 proteinase of hepatitis C virus shows activity in cell-based and in vitro assays. *J Gen Virol 76*:3009–3019.

Perona JJ, Craik CS. 1995. Structural basis of substrate specificity in the serine proteases. *Protein Sci 4*:337–360.

Reed KE, Rice CM. 1998. Molecular characterization of hepatitis C virus. *Curr Stud Hematol Blood Transfus*:1–37.

Sali DL, Ingram R, Wendel M, Gupta D, McNemar C, Tsarbopoulos A, Chen JW, Hong Z, Chase R, Risano C, Zhang R, Yao N, Kwong AD, Ramanathan L, Le HV, Weber PC. 1998. Serine protease of hepatitis C virus expressed in insect cells as the NS3/4A complex. *Biochemistry 37*:3392–3401.

Sambrook J, Fritsch EF, Maniatis T. 1989. In: Nolan C, ed. *Molecular cloning: A laboratory manual*, 2nd ed. Plainview, NY: Cold Springs Harbor Press.

Sharara AI. 1997. Chronic hepatitis C. *South Med J 90*:872–877.

Shoji I, Suzuki T, Chieda S, Sato M, Harada T, Chiba T, Matsuura Y, Miyamura T. 1995. Proteolytic activity of NS3 serine proteinase of hepatitis C virus efficiently expressed in *Escherichia coli*. *Hepatology 22*:1648–1655.

Simmonds P. 1996. Virology of hepatitis C virus. *Clin Ther 18 (Suppl B)*:9–36.

Stadhouders PH, Cooreman MP. 1997. Chronic hepatitis C virus disease: An evaluation of procedures for diagnosis and treatment. *Neth J Med 51*:213–224.

Steinkuhler C, Tomei L, De Francesco R. 1996. In vitro activity of hepatitis C virus protease NS3 purified from recombinant Baculovirus-infected Sf9 cells. *J Biol Chem 271*:6367–6373.

Suzuki T, Sato M, Chieda S, Shoji I, Harada T, Yamakawa Y, Watabe S, Matsuura Y, Miyamura T. 1995. In vivo and in vitro trans-cleavage activity of hepatitis C virus serine proteinase expressed by recombinant baculoviruses. *J Gen Virol 76*:3021–3029.

Takamizawa A, Mori C, Fuke I, Manabe S, Murakami S, Fujita J, Onishi E, Andoh T, Yoshida I, Okayama H. 1991. Structure and organization of the hepatitis C virus genome isolated from human carriers. *J Virol 65*:1105–1113.

Tanji Y, Hijikata M, Satoh S, Kaneko T, Shimotohno K. 1995. Hepatitis C virus-encoded nonstructural protein NS4A has versatile functions in viral protein processing. *J Virol 69*:1575–1581.

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

Tomei L, Failla C, Vitale RL, Bianchi E, De Francesco R. 1996. A central hydrophobic domain of the hepatitis C virus NS4A protein is necessary and sufficient for the activation of the NS3 protease. *J Gen Virol 77*:1065–1070.

Vishnuvardhan D, Kakiuchi N, Urvil PT, Shimotohno K, Kumar PK, Nishikawa S. 1997. Expression of highly active recombinant NS3 protease domain of hepatitis C virus in *E. coli*. *FEBS Lett 402*:209–212.

Yan Y, Li Y, Munshi S, Sardana V, Cole JL, Sardana M, Steinkuehler C, Tomei L, De Francesco R, Kuo LC, Chen Z. 1998. Complex of NS3 protease and NS4A peptide of BK strain hepatitis C virus: A 2.2 Å resolution structure in a hexagonal crystal form. *Protein Sci 7*:837–847.

Zhang R, Durkin J, Windsor WT, McNemar C, Ramanathan L, Le HV. 1997. Probing the substrate specificity of hepatitis C virus NS3 serine protease by using synthetic peptides. *J Virol 71*:6208–6213.