

MOLECULAR DETERMINANTS OF TRIF PROTEOLYSIS MEDIATED BY THE HEPATITIS C VIRUS NS3/4A PROTEASE*

Josephine C. Ferreon¹, Allan Chris M. Ferreon^{2,3}, Kui Li¹, Stanley M. Lemon^{1*}
From the Departments of ¹Microbiology & Immunology, and ²Human Biological Chemistry & Genetics, ¹Center for Hepatitis Research, Institute for Human Infections and Immunity, and ³Sealy Center for Structural Biology, University of Texas Medical Branch at Galveston, Galveston TX 77555-1019

Running title: Molecular Determinants of TRIF Proteolysis by NS3/4A

*Address correspondence to: Stanley M. Lemon, MD, Institute for Human Infections and Immunity, 5.220 Mary Moody Northen Pavilion, University of Texas Medical Branch, 301 University Boulevard, Galveston, TX 77555-0428; Tel: 409-747-6500; Fax: 409-747-6514; E-Mail: smlemon@utmb.edu

Persistent infections with hepatitis C virus (HCV)¹ are a major cause of liver disease and reflect its ability to disrupt virus-induced signaling pathways activating cellular antiviral defenses. HCV evasion of double-stranded RNA signaling through Toll-like receptor 3 is mediated by the viral protease, NS3/4A, which directs proteolysis of its Pro-rich adaptor protein, Toll-IL1 receptor domain containing adaptor inducing interferon- β (TRIF). The TRIF cleavage site has remarkable homology with the viral NS4B/5A substrate, although an 8 residue poly-proline track extends upstream from the P6 position in lieu of the acidic residue present in viral substrates. CD spectroscopy confirmed that a substantial fraction of TRIF exists as poly-proline II helices, and inclusion of the poly-proline track increased affinity of P-side TRIF peptides for the HCV-BK protease. A poly-proline II peptide representing an SH3-binding motif (PPPVPPIRRR, Sos) bound NS3 with moderate affinity, resulting in inhibition of proteolytic activity. Chemical shift perturbations in NMR spectra indicated that Sos binds a 3₁₀ helix close to the protease active site. Thus, a poly-proline II interaction with the 3₁₀ helix likely facilitates NS3/4A recognition of TRIF, indicating a significant difference from NS3/4A recognition of viral substrates. Since SH3-binding motifs are also present in NS5A, a viral protein that interacts with NS3, we speculate that the NS3 3₁₀ helix may be a site of interaction with other viral proteins.

Hepatitis C Virus (HCV) is the causative agent of chronic hepatitis C, a globally distributed

infection that affects more than 170 million persons worldwide, and results in 8,000-10,000 deaths due to liver disease annually in the United States alone (1;2). Currently available therapeutic regimens include combination therapy with interferon and ribavirin, but these are limited in efficacy, frequently associated with adverse reactions, and costly (3). Thus, there is a compelling need for new antiviral drugs possessing greater efficacy against this virus.

HCV is a member of the family *Flaviviridae*, classified within the genus *Hepacivirus*. It has a relatively small, 9.7 kb positive-strand RNA genome, which contains a large open reading frame that spans most of the genomic RNA (4). Translation of genomic RNA results in the expression of a lengthy polyprotein that is co- and post-translationally processed into at least 10 functional proteins by both host and viral protease activities. The processing events that liberate the nonstructural HCV proteins required for viral RNA replication (NS3, NS4A, NS4B, NS5A and NS5B) are directed either *in cis* or *in trans* by a serine protease formed by the noncovalent association of NS3 with a segment of NS4A (5;6). Given the clinical success of inhibitors of the human immunodeficiency virus (HIV) protease (7), the HCV NS3/4A protease has become a leading target for drug discovery efforts. Candidate NS3/4A protease inhibitors have entered clinical trials and have shown substantial promise (8). However, highly active NS3/4A inhibitors have proved to be quite challenging to develop since the active site of the viral protease is unusually featureless (9-11). The shallow contour and solvent-exposed nature of the NS3/4A substrate binding site lacks the pockets or crevices

that would facilitate design of small molecule inhibitors with high affinity and binding specificity.

In addition to its critical role in processing the viral proteins that comprise the viral RNA replicase, the NS3/4A protease disrupts innate intracellular antiviral defenses by blocking virus activation of interferon regulatory factor 3 (IRF-3) and NF- κ B (12;13). These are cellular transcription factors that induce the expression of a large number of cellular antiviral defense genes, including the type 1 α/β interferons and interferon-stimulated genes (ISGs), as well as chemokines and proinflammatory cytokines (14;15). Recent work indicates that the products of viral replication may lead to IRF-3 and NF- κ B activation through two distinct and independent pathways, one involving engagement of Toll-like receptor 3 (TLR3) within endocytic vesicles by viral dsRNA, and the other involving recognition of intracellular structured viral RNAs by the cellular DExH/D RNA helicase, retinoic acid inducible gene I (RIG-I) (16;17). Both pathways are specifically disrupted by the protease activity of the NS3/4A complex (13;18). This disruption of anti-viral signaling can be reversed by specific peptidomimetic inhibitors of the NS3/4A protease, suggesting that it targets for proteolysis one or more proteins involved in these signaling pathways. While, the cellular protein that resides within RIG-I pathway and that is putatively cleaved by NS3/4A has yet to be identified, we have recently demonstrated that NS3/4A targets an essential protein within the TLR3 pathway for proteolysis: Toll-IL1 receptor domain containing adaptor inducing interferon- β (TRIF or TICAM-1) (13). TRIF is an essential adaptor protein that links TLR3 to downstream activation of IRF3 and NF- κ B (19;20), and its cleavage by a viral protease would be expected to disrupt dsRNA-induced signal transduction through the TLR3 pathway. Consistent with this, we have shown that NS3/4A blocks the activation of the interferon- β promoter normally induced by exposure to extracellular poly-I:C, a synthetic dsRNA analog, both in osteosarcoma cells that conditionally express the protease and in HeLa cells supporting on-going replication of subgenomic HCV RNA (13).

TRIF also supports MyD88-independent signaling following the engagement of TLR4 by

pathogen-specific ligands (21). It is thought to interact with the TLRs through a Toll-IL-1 receptor (TIR) homology domain, and to recruit multiple molecular signaling partners through specific domains in its N-terminal and C-terminal sequences (22). The signaling pathways it activates play critical roles in the response of cells to virus infection, and the ability of HCV to disrupt this signaling is likely to contribute substantially to its capacity for sustaining persistent infections in the face of both innate and adaptive immune responses (13). TRIF is cleaved proteolytically by the HCV protease between its Cys₃₇₂ and Ser₃₇₃ residues (13), effectively separating the TIR domain of the protein from an N-terminal TANK-binding kinase 1 (TBK1) interaction site required for IRF-3 phosphorylation, which is a prerequisite for dimerization, nuclear translocation, and activation of IRF-3 as a transcriptional factor (23).

Here, we describe studies of the molecular properties of TRIF that contribute to its ability to function as a substrate for the NS3/4A protease. We demonstrate that the amino acid residues of the protease that interact with TRIF to facilitate its proteolysis differ significantly from those interacting with the canonical viral substrates. The TRIF cleavage site lacks a conserved P6 acidic residue that has been shown in several studies to make a substantial contribution to viral substrate specificity and binding affinity (24;25). This is replaced in TRIF with an 8-residue poly-proline track for which we demonstrate a unique role in interactions with the protease.

EXPERIMENTAL PROCEDURES

Peptides - Synthetic peptides representing sequences spanning the NS3/4A cleavage sites at the viral NS5A/5B (H-EDVV α C/SMSY-OH) (26), NS4A/4B (H-EFDEMEEC/ASHLPYI-OH), and NS4B/5A (H-ECTTPC/SGSWLRD-OH) junctions, the NS3/4A cleavage site within TRIF (Ac-PSSTPC/SAHL-amide, Ac-PPPPPPPSSTPC/SAHL-amide) (13), their P-side cleavage products (Ac-PSSTPC-OH, and Ac-PPPPPPPSSTPC-OH) and an HCV peptide suitable for use in a fluorescence resonance energy transfer (FRET) assay (Ac-DED(EDANS)EEAbu ψ [COO]ASK-(DABCYL)-amide) (27) were purchased from Anaspec and were >95% pure. Sos (Ac-

PPPVPPIRRR-amide) and SosY (Ac-VPPPVPPIRRRY-amide) (28) peptides were synthesized within the UTMB Peptide Synthesis Core Laboratory, and kindly provided by Dr. Vincent Hilser with >95% purity. All peptides, except for SosY, were quantified by amino acid analysis. The concentration of SosY was measured by absorbance at 280 nm using Edelhoch reagent (29).

Protein Over-expression and Purification - TRIF was cloned as a His-tag fusion protein into pET21d vector (Novagen), and recovered from *E. coli* grown in a 20 L fermentor as follows. A single colony of *E. coli*, carrying the expression plasmid, was inoculated into a small flask (37°C, overnight), and then transferred into 20 L LB broth supplemented with 100 µg/mL ampicillin. Cells were grown at 37°C until the OD₆₀₀ reached 0.6-0.8, then induced by addition of 1 mM IPTG followed by further incubation at 37°C for 4-5 hours. Cells were harvested and stored at -80°C prior to purification.

Immunoblots suggested that almost all of the expressed TRIF was present in inclusion bodies; therefore purification was carried under denaturing conditions. The cell pellet was dissolved in extraction buffer (6 M GdnHCl, 100 mM NaPhos, 10 mM Tris, 2 mM β-Me, pH 8, 5 ml of buffer per g of cell pellet) for 24 hours at 4°C, with stirring. Cell debris was removed by ultracentrifugation, and the supernatant passed through a Ni²⁺ affinity FPLC column. The column was rinsed with 10-15 volumes of wash buffer (8 M Urea, 100 mM NaPhos, 10 mM Tris, 2 mM β-MeOH, pH 7.5) containing 50 mM imidazole. Bound proteins were eluted with wash buffer containing 500 mM imidazole. The initial fractions contained TRIF and other low molecular weight contaminants, while the final fractions contained mostly TRIF (>90% purity). Fractions with impurities were subjected to size exclusion chromatography (Superdex 200). TRIF was refolded by dialysis against 20 mM Tris, 2 mM DTT, pH 7.5, with the signature of the folded protein monitored by fluorescence and CD spectroscopy. From a 100 L large-scale purification, approximately 50 mg was obtained with purity of >90%. The majority of the contaminants appeared to be TRIF fragments or degradation products. MALDI-TOF mass

spectrometry (UTMB Protein Facility) confirmed that the molecular mass was approximately 77.2 kDa.

The HCV NS3 single-chain protease (scNS3) and a full-length single chain NS3 (FL scNS3) containing the helicase domain were kindly provided by Bruce Malcolm (Schering-Plough Research Institute); both were derived from the BK strain of HCV (30).

Circular Dichroism (CD) Spectroscopy - Far-UV CD spectra of TRIF were recorded from 200-260 nm using an AVIVTM CD spectrometer with the temperature maintained at 25°C and the bandwidth set at 1 nm with a 0.5 sec averaging time. The protein concentration was approximately 0.4 mg/ml in 20 mM Tris, 2 mM DTT, pH 7.5. Buffer spectra were collected and subtracted from the protein spectra.

Comparative Kinetic Analyses of NS3 Protease Activity with Different Viral and TRIF Substrates - In vitro cleavage of the different peptide substrates (5A/5B, 4A/4B, 4B/5A, TRIF p372, and HCV-FRET) was monitored using an HPLC ShimadzuTM chromatograph, equipped with UV-VIS detection. For the 5A/5B, 4A/4B, 4B/5A substrates, the peptide fragments were monitored at 280 nm wavelength, versus 220 nm for TRIF p372 and 512 nm for the FRET peptide. The reaction mix consisted of varying concentrations of the peptide substrate and enzyme (scNS3, BK strain) in 25 mM Tris, 300 mM NaCl, 5 µM EDTA, 10% glycerol, 0.05% n-dodecyl-β-D-maltoside, 10 mM DTT, pH 7.5. Enzyme concentrations were 4.2 nM for the 5A/5B substrate, 85.7 nM for 4A/4B, 1 µM for 4B/5A, and 3.3 µM for TRIF p372 substrate, with substrate concentrations ranging from 15-70 µM for 5A/5B, 70-700 µM for 4A/4B, 100-1800 µM for 4B/5A, and 100-3500 µM for the TRIF p372 substrate. For most of the peptide substrates, such as 4B/5A and TRIF p372, the final peptide concentrations did not exceed the K_m due to poor solubility. Nonetheless, enzyme concentrations generally were at least 100-fold less than substrate concentrations. Preliminary experiments ensured linearity in the initial velocity under these conditions, and demonstrated that the enzyme concentration was directly proportional to the initial velocity. Reactions were incubated at 30°C

for a period of time sufficient to achieve ~20% cleavage, then quenched with 1% TFA. The products analyzed by HPLC, with the fragments separated using a 10-50% acetonitrile gradient at 2.5 ml/min. Areas of product peaks were integrated and quantitated with calibration standards. Kinetic parameters (k_{cat} and K_m) were obtained by non-linear fitting of the data (initial rates versus substrate concentration) with the Michaelis-Menten equation, $V_0 = k_{cat}[E_t][S]/([S] + K_m)$, where V_0 is the initial velocity, E_t is the total enzyme concentration, S is the substrate concentration, k_{cat} is the catalytic turnover and K_m is the Michaelis constant.

Full-length TRIF Cleavage by Full-length NS3 and scNS3 - For proteolytic cleavage of TRIF by the full-length NS3 (both protease and helicase domains), reactions contained 4 μ M TRIF (20 mM Tris, 30 mM DTT, pH 7.5) and 2 μ M NS3 or scNS3 (BK strain), with or without 10 μ M SCH6, a ketoamide, peptidomimetic NS3/4A protease inhibitor (12). Reactions were incubated at 28°C for 2 hrs.

Estimated EC₅₀ for scNS3 Cleavage of Full-length TRIF - To determine the enzyme concentration required to hydrolyze 50% of the full-length TRIF protein (EC₅₀), different reaction tubes were prepared with 4 μ M TRIF and varying scNS3 concentrations (0.25, 0.5, 1, 2, 3, 5 μ M) in 20 mM Tris, 30 mM DTT, pH 7.5. Reactions were incubated at 30°C for 30 min, quenched with gel loading buffer, and frozen at -20°C prior to SDS-PAGE analysis. The quantity of the remaining full-length TRIF species present in each reaction mix was estimated by densitometric analysis. The data (TRIF intensity vs. enzyme concentration) was fitted in a non-linear fashion to the following simplified four-parameter logistic function (27) derived from the Hill equation: $y = 100/(1 + ([E]/EC_{50})^n)$; where y is the substrate intensity (TRIF), normalized to 100% for no cleavage, E is the enzyme concentration, and n is the Hill slope.

Estimated EC₅₀ for scNS3 Cleavage of the TRIF p372 peptide - Reaction mixtures contained 4 μ M of TRIF p372 and varying scNS3 concentrations: 0, 1, 2, 5, 8.3, and 16.7 μ M. The buffer conditions were identical to those for the full-length TRIF cleavage assay. The reactions were incubated at

30°C for 30 min and quenched with an equal volume of 1% TFA, prior to HPLC analysis. The decrease in the quantity of p372 was plotted against the enzyme concentration and fitted to the equation described in the preceding section.

Fluorescence Spectroscopy - Binding of the scNS3 protease to the Sos peptide, and a short and long TRIF peptide (Ac-PSSTPC-OH and Ac-PPPPPPPSSTPC-OH, respectively) were followed by fluorescence spectroscopy (SPEX FluoroMaxTM spectrofluorometer). The excitation and emission wavelengths were fixed at 275 and 340 nm, respectively, to monitor tyrosine and tryptophan fluorescence changes. The observed decreases in the fluorescence intensity signal that occurred upon ligand binding were fitted to a single-site binding model equation: $y = A*x/(K_d + x)$ where y is the relative fluorescence intensity, x is the ligand concentration, A is the maximum fluorescence intensity and K_d is the dissociation constant. The experiments were performed at 25°C and in 75 mM Potassium Phosphate buffer, 10 mM DTT, pH 6.5.

IC₅₀ Determination of Sos Peptide Inhibition of NS3/4A Protease Activity - Protease activity was assayed using the FRET-based HCV peptide as substrate, monitored by HPLC at 512 nm. The concentration of the SosY peptide necessary to inhibit 50% of the reaction (IC₅₀) was determined from the decrease in the initial velocity of the reaction observed with increasing concentrations of the SosY peptide, fitted to the equation $y = A/(1 + ([I]/IC_{50})^n)$, where y is the initial velocity, A is the maximum initial velocity, I is the inhibitor concentration, and n is the Hill slope. The buffer conditions were 50 mM Tris, 50% glycerol, 2% CHAPS, 30 mM DTT, pH 7.5.

NMR Spectroscopy - NMR spectra were collected at 25 °C using a Varian UnityPlus 750 MHz instrument, equipped with a triple resonance probe and a pulsed field gradient. Sensitivity-enhanced ¹H-¹⁵N- HSQC spectra were recorded with identical acquisition parameters and 0.5 mM ¹⁵N-labeled scNS3 protein (75 mM Potassium phosphate, 5% glycerol, 25 mM DTT, 0.015% NaN₃, 10% D₂O, pH 6.5) with different Sos peptide concentrations (0, 0.3, 1.3, 2.2 mM final). Data were processed and visualized on an SGI workstation using Felix 98.0 software. To increase

spectral resolution, time domain data were zero-filled twice and 90° phase-shifted sinebell apodization functions were applied in both dimensions. The protein and chemical shift assignments (31) were kindly provided by Bruce Malcolm (Schering Plough Research Institute).

Vizualizations - The Molmol program (32) was used to generate ribbon models of the NS3 protease (PDB entry 1NS3) (Fig. 7), and the ball-and-stick model of the Sos peptide (Fig. 4B), adopted from the co-crystal structure of the SEM-5 C-terminal SH3 domain with the Sos peptide (PDB entry 1SEM, chain C) (33).

RESULTS

TRIF is a Novel Host Substrate for the HCV NS3/4A Protease. NS3 is a bifunctional protein, with a serine protease domain located within its amino terminal third, and ATPase/RNA helicase activity located within its carboxy terminal two-thirds (34). The NS3/4A protease is responsible for directing cleavage within the HCV polyprotein at the NS3/4A, 4A/4B, 4B/5A, 5A/5B junctions. NS4A remains noncovalently associated with NS3 following scission at the NS3-4A junction, and it is a necessary cofactor for the full expression of NS3 protease activity (6;35;36). In the studies presented here, we used a single-chain protease, scNS3, in which residues 21-32 of NS4A are fused to the amino terminus of the protease domain of NS3 (30;37). This construction allows the NS4A cofactor peptide sequence to fold properly to form a mature NS3/4A protease possessing activity that is very similar if not identical to that of the noncovalent NS3/4A complex. In previous demonstrations of TRIF cleavage by NS3/4A, we used both an scNS3 protein and a bacterially-expressed product representing the NS3 protease domain complexed with an NS4A peptide cofactor; both derived from the HCV-N strain of HCV (13). In the experiments described here, we used proteases expressed from the HCV-BK strain, since we encountered less aggregation at higher concentrations of these proteins. Use of a protease from an additional HCV strain also allowed us to demonstrate that TRIF hydrolysis is not limited to a single strain of HCV.

Given the bifunctional nature of NS3/4A, we first set out to establish whether there is any

difference between the ability of a scNS3 protein representing only the protease domain of NS3, and a full-length scNS3 (FL scNS3) containing both the protease and helicase domains, to cleave TRIF. As shown in Fig. 1A, purified bacterially-expressed TRIF was cleaved at the same position both by scNS3 (protease domain only) and the FL scNS3 (complete NS3 protein containing both protease and helicase domains, lanes 2 and 6, respectively). These reactions yielded two fragments, Δ C340 and Δ N372, representing the N-terminal (P-side) and C-terminal (P'-side) fragments respectively (Fig. 1A) (13). The addition of SCH6, a specific peptidomimetic ketoamide inhibitor of the HCV protease, abolished the cleavage (Fig. 1A, lanes 3 and 7). These results confirm that TRIF proteolysis is mediated by the NS3/4A protease, as shown before (13), regardless of the strain or type of construct.

To further characterize TRIF proteolysis by the viral protease, a 10-residue peptide substrate (spanning residues P₆-P'₄) was designed based on the amino acid sequence of TRIF surrounding the NS3/4A cleavage site (TRIF p372, PSSTPCSAHL). This length is the minimum generally required for efficient NS3/4A proteolysis of viral substrates, and includes the acidic residue conserved at the P6 position of the viral polyprotein *trans*-cleavage substrate sites. This acidic residue contributes to catalytic efficiency through electrostatic interactions with the protease (24;25). On the other hand, lengthier peptides representing the viral substrate generally do not show appreciably greater substrate activity. Fig. 1B shows a three-dimensional representation of an HPLC chromatogram of the peptide cleavage reaction where the substrate 10-mer peptide (TRIF p372, PSSTPCSAHL) decreases in abundance in concert with increases in the peptide products as a function of reaction time. The products increase in a hyperbolic fashion characteristic of a Michaelis-Menten enzymatic reaction.

Kinetics of Proteolysis of TRIF versus Viral Substrates. The sequence of TRIF p372 on the P-side of the Cys-Ser cleavage site is quite similar to that at the NS4B/5A junction, except for the notable difference in the P6 position which is a conserved acidic residue in the viral substrate, as noted above (Table 1, shaded) (13). The role of

these conserved, negatively-charged residues in the viral cleavage sites has been investigated by Koch et al. (24), who have shown that peptides representing the NS5A/5B cleavage site with substitutions resulting in neutral charges at the P6 and P5 position possess as much as a 55-fold difference in K_m while exhibiting similar k_{cat} values. The P'-side of TRIF p372 is quite similar to the NS4A/4B site, except that the Ser and Ala residues at the P'1 and P'2 positions are reversed (Table 1).

We compared the differences in proteolysis between the TRIF peptide and peptide substrates based on the viral cleavage sequences using the HCV-BK scNS3 protease in HPLC-based cleavage assays (see Methods for details). Table 2 summarizes the kinetic parameters for viral substrates based on the NS5A/5B, NS4A/4B, NS4B/5A junctions within the viral polyprotein (Table 1), as well as for the cellular substrate TRIF p372 peptide. The affinity of the protease for the TRIF p372 peptide was relatively weak ($K_m \sim 10$ mM), with a k_{cat} of 14.5 min^{-1} and k_{cat}/K_m of $23 \text{ M}^{-1} \text{ s}^{-1}$ (Table 2). Proteolysis of the TRIF p372 peptide proceeded more slowly than with the viral substrate peptides, with the following rank order for cleavage efficiency: NS5A/5B > NS4A/4B > NS4B/5A > TRIF p372 (see Table 2). The rank order of the viral substrates was similar to that observed in previous studies (10). From Table 2, it can be seen that the difference in proteolysis of the TRIF p372 and viral substrates, 4A/4B and 4B/5A, is due largely to differences in K_m (typically interpreted as the affinity for the substrate). The k_{cat} were comparable and within the range of error in the assay. Nonetheless, the k_{cat}/K_m for the TRIF p372 peptide was approximately half that observed for the most slowly cleaved 4B/5A peptide.

Kinetics of Cleavage of Full-length TRIF Versus the TRIF Peptide. From the kinetic studies above, we concluded that proteolysis of the TRIF peptide is less efficient than the viral substrates, although still within the range of what might be expected for a biologically relevant cleavage. One possible explanation for the reduced cleavage observed with the TRIF peptide is the absence of an acidic residue at the P6 position that has been shown to contribute to enhancement in K_m values with viral substrates. In the TRIF sequence, the P6 residue is proline and it is preceded by a string of 7 other

proline residues. Poly-proline sequences are ubiquitous in cell signaling molecules and they have been shown to play a role in the regulation of cellular processes by binding to SH3, WW, and EVH domains (38;39). Consistent with its role as a signaling molecule, TRIF is highly proline-rich and we reasoned that this might play a role in its recognition by the HCV protease.

To test the hypothesis that there might be other determinants beyond the 10 residues flanking the cleavage site that could enhance its affinity for NS3/4A and facilitate subsequent proteolysis, we compared the kinetics of scNS3-mediated cleavage of full-length, bacterially-expressed TRIF protein and the TRIF p372 peptide. Since it was difficult to achieve a concentration of the full-length TRIF sufficient for determination of the k_{cat}/K_m of this cleavage reaction, we estimated the difference in the enzyme concentrations (EC_{50}) required to achieve 50% cleavage of the peptide and full-length protein from the Hill equation, keeping other conditions constant (see Methods for details). Fig. 2 shows that there was nearly complete cleavage of the full-length TRIF substrate at a scNS3 concentration of $5 \mu\text{M}$ (Fig. 2A), while only 10% of the peptide substrate was hydrolyzed at this enzyme concentration (Fig. 2B). A quantitative analysis suggested that the EC_{50} for full-length TRIF is $0.69 \pm 0.04 \mu\text{M}$, compared with $27 \pm 3 \mu\text{M}$ for TRIF p372 (Fig. 2C), indicating a 40-fold difference in the rate of proteolysis. These results thus clearly indicate a substantial difference between the two substrates, leading us to conclude that there are other interaction sites beyond the p372 peptide sequence that contribute to the recognition of TRIF by NS3/4A and that enhance the efficiency of proteolysis.

TRIF is a Pro-rich Protein. As indicated previously, the TRIF sequence contains abundant proline residues (13.6%, 97 out of 712 residues). This is consistent with its aberrant migration in SDS-PAGE, where it appears to have an apparent molecular mass of >94 kDa under denaturing conditions (Fig. 1A and 2A), even though mass spectral analysis confirmed that the purified recombinant TRIF possessed the expected mass of 77.2 kDa. This anomalous migration is most likely attributable to poly-proline kinks in the molecule, as observed in other Pro-rich proteins (40).

Importantly, the NS3/4A cleavage of TRIF between Cys₃₇₂ and Ser₃₇₃ (as demonstrated previously) (13) should result in 2 fragments (Δ C340 and Δ N372) of almost equal molecular mass (39 and 38 kDa for the N-terminal and C-terminal fragments, respectively). However, the two cleavage products are clearly separated in SDS-PAGE, with the N-terminal fragment (which contains 59 of the 96 proline residues) showing the characteristic anomalous migration (>43 kDa, see Fig. 1A and 2A). There are numerous poly-proline stretches in the protein, including notably, the 8 residue poly-proline track within the N-terminal cleavage fragment, just proximal to the cleavage site (Fig. 3A).

Pro-rich sequences are recognized by a wide variety of signaling molecules that contain SH3, WW, EVH1 and GYF domains (see (38;39) and references therein). Typical signaling molecules contain more than one recognition module. An example is the cellular Grb2 protein that has 2 SH3 domains and 1 SH2 domain. SH3 domains recognize Pro-rich motifs with moderate to weak affinities, usually in the μ M to mM range (39;41). These interactions, although not very strong, involve specific hydrogen-bond interactions and rigid intercalation of aromatic side chains (33). As illustrated by the Grb2 protein, these Pro-rich sequence-binding motifs are generally part of larger recognition domains that are capable of interacting with multiple domains of the binding partner, such that the weak interactions result in enhanced specificity and affinity. A characteristic feature of these Pro-rich sequences is that they form a left-handed poly-proline II (PPII) helix, which is characterized by proline residues in the *trans* position. To determine whether TRIF possesses such structure, we analyzed the CD spectra of the purified full-length TRIF. As shown in Fig. 3B, we observed a substantial dip in the ellipticity at \sim 205 nm, which is indicative of the presence of PPII helices (42;43).

These observations led us to speculate that the poly-proline stretch just upstream of the cleavage site could play an important role in enhancing the affinity of TRIF for the NS3 protease. Based on knowledge of the structural basis of the interaction between SH3 domains and Pro-rich polypeptides (33;44), we looked for a potential binding site within the NS3/4A protease.

Interestingly, we identified a 3_{10} helix, composed of hydrophobic residues (Ile-132, Tyr-134, Leu-135), in close proximity to the protease active site. We reasoned that this could act as a possible site for interaction with the poly-proline track in TRIF.

The NS5A protein is a nonstructural protein expressed by HCV that has no precise function identified as yet in viral RNA replication (4), but appears to contribute to viral disruption of cellular antiviral defenses through an interaction with protein kinase R (for a review, see (45) and references therein). NS5A is also a Pro-rich protein, containing 11% prolines and encoding two PXXP motifs such as are capable of interacting with SH3 domains (46). In the genotype 1a H77c strain of HCV, one of these Pro-rich motifs has a sequence that is identical to the PXXP motif within the murine Son of sevenless (Sos) homolog protein (PPVPPRR, see Fig. 4B), while in the genotype 1b BK strain this sequence is closely homologous (possessing a substitution of the central hydrophobic Val residue in Sos with Ile). The peptide representing the Sos domain (PPPVPPRRR, Fig. 4A) has been shown to interact with the SH3 domain of Grb2 involved in the Ras-Erk pathway (47;48). Not surprisingly, therefore, the viral NS5A protein has also been shown to interact with Grb2 and perturb mitogenic signaling (46;49). The Sos peptide is a model PPII helix (Fig. 4B), and thus potentially representative of many of the Pro-rich stretches in TRIF.

Sos Peptide Interacts with scNS3. To determine whether the Pro-rich Sos peptide is capable of interacting specifically with the scNS3 molecule, we assessed binding using fluorescence spectroscopy and demonstrated a signal decrease in fluorescence intensity upon addition of the peptide. Fig. 5A shows a normalized plot of the change in fluorescence intensity as a function of Sos peptide concentration. The data were fitted to a single-site binding model equation (see Methods), resulting in a dissociation constant of 0.53 ± 0.13 mM. The affinity of Sos for scNS3, albeit within the μ M range, was thus measurable and characteristic of proline interactions.

If the Sos peptide does interact with the protease near the active site, such as at the 3_{10} helix, it could compete with substrate for binding to the scNS3 protease. We assessed this possibility using a FRET-based peptide as substrate, and

detecting proteolysis by HPLC at 512 nm. The concentration of a SosY peptide (VPPPVPPIRRRY) necessary to inhibit 50% of the cleavage reaction (IC_{50}) was determined from the decrease in the initial velocity of the reaction as a function of increasing SosY peptide concentration, and fitted to the equation: $y = A/(1 + ([I]/IC_{50})^n)$ (Fig. 5B). These results indicated that the SosY peptide is capable of inhibiting the scNS3 protease activity with an $IC_{50} = 2.7 \pm 0.2$ mM. The derived value of $n = 1.3 \pm 0.1$ is consistent with single-site binding. Inhibition was also observed in a similar experiment carried out with scNS3 derived from the HCV-N strain of HCV (data not shown).

Localization of the Sos Peptide Interaction with NS3/4A. To determine the specific site of interaction of the Sos peptide with scNS3, we used nuclear magnetic resonance (NMR) to determine chemical shift changes occurring in a purified ^{15}N -labeled single-chain protease in response to the binding of the peptide. Fig. 6A shows representative segments of the 1H - ^{15}N HSQC spectra of the unbound (red) and bound (blue) protease. The residues that comprise the catalytic triad that are responsible for the proteolytic reaction are H57, D81 and S139. Upon binding of the Sos peptide, there was a significant chemical shift change in the S139 residue (Fig. 6A, left panel), while in contrast R123, for example, did not demonstrate a shift (Fig. 6A, right panel). Fig. 6B shows the magnitude of the chemical shift differences between the unbound and bound protease at each residue. The greatest differences occurred within residues in the 3_{10} helix, parts of β -strand E2 ($\beta E2$), and residues contributing to the protease active site (Fig. 7). These results are distinctly different from previous reports of NMR studies of viral substrate-based inhibitors (50). Previous studies with DEDifEChC-OH (an optimized peptidomimetic inhibitor based on the NS4A/4B P-side product) suggest that this peptide also interacts with $\beta E2$, forming extensive H-bond interactions. However, the DEDifEChC-OH binding site also included the $\beta F2$ strand and certain residues in $\beta C2$. Contrary to what we observed for Sos peptide binding, its binding was associated with large shifts in residues in $\beta C2$ and $\beta F2$. In the case of Sos, while many scNS3 residues, including S139, S138, K136, A157, H57 demonstrated significant chemical shift changes

(Fig. 6B), most of the residues contributing to $\beta C2$ and $\beta F2$ did not shift significantly (e.g., R123, A45, D168) (Fig. 6A). These data thus suggest that Sos binds primarily to the 3_{10} helix side, possibly affecting some residues in $\beta E2$ due to proximity. While the Sos peptide lacks a reactive Cys, it is nonetheless capable of perturbing the catalytic site residues, S139, H57, D81 (Fig. 6B), which is consistent with both the binding and inhibition data (Fig. 6). Since the sequence of the Sos SH3-binding domain is present within NS5A (Fig. 4A), the demonstration of a Sos-interaction domain on NS3 raises the possibility that NS3 may interact with NS5A in a similar fashion during assembly of the viral RNA replicase (see Discussion).

In the absence of the lengthy poly-proline track present in TRIF, the P6-P1 residues of the TRIF p372 peptide (PSSTPCSAHL) are likely to bind to the $\beta E2$ strand of NS3 in a manner similar to that observed in studies of viral substrate-based product inhibitors. However, the NMR data suggest that the upstream poly-proline segment in TRIF (Fig. 3A), by interacting with the 3_{10} helix, may enhance the affinity of TRIF for the NS3/4A protease, possibly explaining the enhanced proteolytic rate we observed for the full-length TRIF compared with the TRIF p372 peptide (Fig. 2). To assess the impact of the poly-proline sequence on cleavage of a peptide substrate, we compared the rates of cleavage of the TRIF p372 peptide with a peptide containing the upstream poly-proline sequence (PPPPPPPSSTPCSAHL) (Fig. 8A). The low solubility of this extended peptide precluded a complete determination of the cleavage kinetics. However, at equivalent substrate concentrations, the proteolytic cleavage of the extended peptide substrate proceeded at a slower rate than the TRIF p372 peptide (Fig. 8A). This was not unexpected, as the P-side product of the cleavage of the extended peptide is likely to dissociate more slowly from NS3 following cleavage than the TRIF p372 peptide P-side product, and to possess significant protease inhibitory activity as shown for the Sos peptide in Fig. 5B.

To directly assess this poly-proline 'anchoring' hypothesis we used fluorescence spectroscopy to determine the difference in the NS3-binding affinity of short and long TRIF P-

side oligopeptides that either did or did not include the upstream poly-proline segment. These experiments demonstrated that the extended TRIF peptide containing the upstream poly-proline segment possessed approximately 10-fold greater affinity for the protease than a short peptide lacking the poly-proline sequence (0.15 ± 0.03 and 1.2 ± 0.1 mM for PPPPPPPSSTPC and PSSTPC TRIF peptides, respectively) (Fig. 8B).

DISCUSSION

We have shown here that there is a significant difference (~40-fold) in the rate of NS3 cleavage of full-length TRIF and the short TRIF p372 peptide. It is likely that this difference is due to features of the TRIF molecule that direct its binding to NS3 outside of the protease active site, including the poly-proline track immediately upstream of the NS3 cleavage at Cys₃₇₂ of TRIF. The NMR studies (Fig. 6) and increased affinity of TRIF peptides containing the poly-proline track (Fig. 8B) strongly suggest that it contributes to recognition of TRIF by NS3. However, it is likely that other features of the full-length TRIF molecule also contribute to its relatively rapid cleavage by the protease. Even with the enhanced affinity of the full-length protein, the rate of TRIF hydrolysis is still far slower than that of the NS5A/5B peptide which has an EC₅₀ in the low nM (<10 nM) (34) range, compared to ~600 nM for TRIF (Fig. 2). TRIF cleavage may be faster than the NS4B/5A cleavage, however, and comparable to the NS4A/4B cleavage. These kinetic studies thus provide additional evidence supporting the potential biological relevance of the NS3/4A-directed cleavage of TRIF in the disruption of cellular antiviral defenses (13). TRIF cleavage could thus play an important role in promoting viral persistence.

PPII motifs are known to play an essential role in cellular signaling (38;39). Furthermore, these motifs may also be involved in viral regulatory processes. For instance, the HIV-1 Nef protein contains Pro-rich motifs that bind to the Src-family SH3 domain promoting enhanced replication of the virus (51). There are several potential mechanisms by which proline-binding domains may enhance the specificity of intermolecular interactions (39). These include the extension of the specificity surface beyond the

active site, as well as the provision of a separate, independent recognition surface. Both seem likely for the TRIF-NS3/4A interaction. The interaction with the protease active site appears to be effectively extended to include the 3₁₀ helix of the protease, which intercalates with the PPII helix of TRIF. Thus, the lack of a P6 acidic residue in TRIF may be compensated by the presence of a 'PPII anchor' that extends beyond the substrate site, resulting in enhanced affinity and specificity. Other PPII segments in TRIF could also interact with the protease at sites remote from the catalytic triad, a possibility that is not excluded by our NMR studies.

Previous studies have shown that the non-structural proteins spanning the NS3-NS5B segment of the HCV polyprotein are required for RNA replication and are likely to contribute to the assembly of a large, macromolecular viral replicase (52;53). Indeed, yeast two-hybrid, GST pull-down, and coimmunoprecipitation assays suggest that NS3 interacts with NS5A (54), and may influence the phosphorylation status of NS5A (55). Such interactions might be moderate to weak, yet still necessary for the transient and dynamic interactions that are likely to be involved in assembly of the macromolecular HCV RNA replicase complex. The presence of the Sos peptide sequence within NS5A (Fig. 4), coupled with demonstration of the affinity of Sos for NS3 (Fig. 5), suggests the possibility that the Sos domain of NS5A could similarly interact with the 3₁₀ helix of NS3. If so, the putative poly-proline interaction between NS5A and NS3 could be strengthened by the interaction of NS3 with its NS4A cofactor sequence, as NS4A also interacts with NS5A (45;56) (see Fig. 4A). There may also be additional, yet uncharacterized interaction sites elsewhere.

Our findings may have significance for drug discovery efforts. Since the HCV NS3 substrate binding site is rather shallow and featureless, the design of specific small molecule inhibitors of this enzyme has been difficult and challenging. The identification of a novel poly-proline binding site that is capable of 'anchoring' peptides near the protease active site might open new avenues for structure-based drug design by targeting both sites for binding interactions. Previous structure-based drug design studies have provided proof of concept for bivalent peptides.

The 'tethering' of two different peptides that bind in weak mM range at closely proximate sites on a target molecule may result in binding affinities in the sub-micromolar range due to the reduction in the rotational and translational entropy associated with the 'tethering' (57). We hypothesize that the TRIF sequence thus acts as a naturally-occurring bivalent ligand, wherein the poly-proline sequence acts as an 'anchor' and further enhances affinity of

the substrate for the protease. In addition, we speculate that small molecule therapeutics that are designed to target both interactions may not only inhibit protease activity, but also potentially interrupt assembly of the HCV replicase complex by disrupting NS3-NS5A binding mediated by the Sos sequence in NS5A. Such molecules might have particularly potent antiviral activity.

REFERENCES

1. Wong, J. B., McQuillan, G. M., McHutchison, J. G., and Poynard, T. (2000) *Am.J Public Health* **90**, 1562-1569
2. Alter, M. J., Mast, E. E., Moyer, L. A., and Margolis, H. S. (1998) *Infect.Dis.Clin.North Am.* **12**, 13-26
3. McHutchison, J. G. and Fried, M. W. (2003) *Clin.Liver Dis.* **7**, 149-161
4. Reed, K. E. F. and Rice, C. M. (2000) *Curr Top Microbiol Immunol* **242**, 55-84
5. Grakoui, A., McCourt, D. W., Wychowski, C., Feinstone, S. M., and Rice, C. M. (1993) *J Virol* **67**, 2832-2843
6. Bartenschlager, R., Lohmann, V., Wilkinson, T., and Koch, J. O. (1995) *J.Virol.* **69**, 7519-7528
7. Randolph, J. T. and DeGoey, D. A. (2004) *Curr Top Med.Chem.* **4**, 1079-1095
8. Lamarre, D., Anderson, P. C., Bailey, M., Beaulieu, P., Bolger, G., Bonneau, P., Bos, M., Cameron, D. R., Cartier, M., Cordingley, M. G., Faucher, A. M., Goudreau, N., Kawai, S. H., Kukolj, G., Lagace, L., LaPlante, S. R., Narjes, H., Poupart, M. A., Rancourt, J., Sentjens, R. E., St George, R., Simoneau, B., Steinmann, G., Thibeault, D., Tsantrizos, Y. S., Weldon, S. M., Yong, C. L., and Llinas-Brunet, M. (2003) *Nature* **426**, 186-189
9. Love, R. A., Parge, H. E., Wickersham, J. A., Hostomsky, Z., Habuka, N., Moomraw, E. W., Adachi, T., and Hostomska, Z. (1996) *Cell* **87**, 331-342
10. Kim, J. L., Morgenstern, K. A., Lin, C., Fox, T., Dwyer, M. D., Landro, J. A., 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
11. De Francesco, R. and Steinkuhler, C. (2000) *Curr.Top.Microbiol.Immunol.* **242**, 149-169
12. Foy, E., Li, K., Wang, C., Sumter, R., Ikeda, M., Lemon, S. M., and Gale, M., Jr. (2003) *Science* **300**, 1145-1148
13. Li, K., Foy, E., Ferreon, J. C., Nakamura, M., Ferreon, A. C. M., Ikeda, M., Ray, S. C., Gale, M. Jr., and Lemon, S. M. (2005) *Proc. Nat'l. Acad. Sci. USA* **102**, 2992-7
14. Hiscott, J., Pitha, P., Genin, P., Nguyen, H., Heylbroeck, C., Mamane, Y., Algarte, M., and Lin, R. (1999) *J Interferon Cytokine Res.* **19**, 1-13
15. Santoro, M. G., Rossi, A., and Amici, C. (2003) *EMBO J.* **22**, 2552-2560
16. Yoneyama, M., Kikuchi, M., Natsukawa, T., Shinobu, N., Imaizumi, T., Miyagishi, M., Taira, K., Akira, S., and Fujita, T. (2004) *Nat.Immunol.* **5**, 730-737
17. Alexopoulou, L., Holt, A. C., Medzhitov, R., and Flavell, R. A. (2001) *Nature* **413**, 732-738
18. Foy, E., Li, K., Sumpter, R., Jr., Loo, M. Y., Johnson, C., Wang, C., Fish, P., Yoneyama, M., Fujita, T., Lemon, S. M., and Gale, M., Jr. (2005) *Proc.Natl.Acad.Sci.USA* **102**, 2986-91
19. Oshiumi, H., Matsumoto, M., Funami, K., Akazawa, T., and Seya, T. (2003) *Nat.Immunol.* **4**, 161-167
20. Sato, S., Sugiyama, M., Yamamoto, M., Watanabe, Y., Kawai, T., Takeda, K., and Akira, S. (2003) *J Immunol.* **171**, 4304-4310
21. Fitzgerald, K. A., Rowe, D. C., Barnes, B. J., Caffrey, D. R., Visintin, A., Latz, E., Monks, B., Pitha, P. M., and Golenbock, D. T. (2003) *J Exp.Med.* **198**, 1043-1055
22. Akira, S. (2003) *J.Biol.Chem.* **278**, 38105-38108
23. Servant, M., Grandvaux, N., and Hiscott, J. (2002) *Biochem.Pharmacol.* **64**, 985
24. Koch, U., Biasiol, G., Brunetti, M., Fattori, D., Pallaoro, M., and Steinkuhler, C. (2001) *Biochemistry* **40**, 631-640
25. Urbani, A., Bianchi, E., Narjes, F., Tramontano, A., De Francesco, R., Steinkuhler, C., and Pessi, A. (1997) *J.Biol.Chem.* **272**, 9204-9209
26. Landro, J. A., Raybuck, S. A., Luong, Y. P., O'Malley, E. T., Harbeson, S. L., Morgenstern, K. A., Rao, G., and Livingston, D. J. (1997) *Biochemistry* **36**, 9340-9348

27. Taliani, M., Bianchi, E., Narjes, F., Fossatelli, M., Urbani, A., Steinkuhler, C., De Francesco, R., and Pessi, A. (1996) *Anal.Biochem.* **240**, 60-67
28. Ferreon, J. C. and Hilser, V. J. (2003) *Protein Sci.* **12**, 447-457
29. Edelhoch, H. (1967) *Biochemistry* **6**, 1948-1954
30. Howe, A. Y., Chase, R., Taremi, S. S., Risano, C., Beyer, B., Malcolm, B., and Lau, J. Y. (1999) *Protein Sci.* **8**, 1332-1341
31. McCoy, M. A., Senior, M. M., Gesell, J. J., Ramanathan, L., and Wyss, D. F. (2001) *J Mol.Biol.* **305**, 1099-1110
32. Koradi, R., Billeter, M., and Wuthrich, K. (1996) *J Mol.Graph.* **14**, 51-32
33. Lim, W. A., Richards, F. M., and Fox, R. O. (1994) *Nature* **372**, 375-379
34. Yao, N., Reichert, P., Taremi, S. S., Prosis, W. W., and Weber, P. C. (1999) *Structure.Fold.Des.* **7**, 1353-1363
35. Gallinari, P., Paolini, C., Brennan, D., Nardi, C., Steinkuhler, C., and De Francesco, R. (1999) *Biochemistry* **38**, 5620-5632
36. Lin, C., Thomson, J. A., and Rice, C. M. (1995) *J.Virol.* **69**, 4373-4380
37. Taremi, S. S., Beyer, B., Maher, M., Yao, N., Prosis, W., Weber, P. C., and Malcolm, B. A. (1998) *Protein Sci.* **7**, 2143-2149
38. Kay, B. K., Williamson, M. P., and Sudol, M. (2000) *FASEB J* **14**, 231-241
39. Zarrinpar, A., Bhattacharyya, R. P., and Lim, W. A. (2003) *Sci.STKE.*
40. Proft, T., Hilbert, H., Plagens, H., and Herrmann, R. (1996) *Gene* **171**, 79-82
41. Ferreon, J. C. and Hilser, V. J. (2004) *Biochemistry* **43**, 7787-7797
42. Kelly, M. A., Chellgren, B. W., Rucker, A. L., Troutman, J. M., Fried, M. G., Miller, A. F., and Creamer, T. P. (2001) *Biochemistry* **40**, 14376-14383
43. Ma, K., Kan, L., and Wang, K. (2001) *Biochemistry* **40**, 3427-3438
44. Ferreon, J. C. and Hilser, V. J. (2003) *Protein Sci.* **12**, 982-996
45. Tan, S. L. and Katze, M. G. (2001) *Virology* **284**, 1-12
46. Tan, S. L., Nakao, H., He, Y., Vijaysri, S., Neddermann, P., Jacobs, B. L., Mayer, B. J., and Katze, M. G. (1999) *Proc.Natl.Acad.Sci.U.S.A* **96**, 5533-5538
47. Gale, N. W., Kaplan, S., Lowenstein, E. J., Schlessinger, J., and Bar-Sagi, D. (1993) *Nature* **363**, 88-92
48. Lowenstein, E. J., Daly, R. J., Batzer, A. G., Li, W., Margolis, B., Lammers, R., Ullrich, A., Skolnik, E. Y., Bar-Sagi, D., and Schlessinger, J. (1992) *Cell* **70**, 431-442
49. He, Y., Nakao, H., Tan, S. L., Polyak, S. J., Neddermann, P., Vijaysri, S., Jacobs, B. L., and Katze, M. G. (2002) *J Virol.* **76**, 9207-9217
50. Cicero, D. O., Barbato, G., Koch, U., Ingallinella, P., Bianchi, E., Nardi, M. C., Steinkuhler, C., Cortese, R., Matassa, V., De Francesco, R., Pessi, A., and Bazzo, R. (1999) *J.Mol.Biol.* **289**, 385-396
51. Lee, C. H., Saksela, K., Mirza, U. A., Chait, B. T., and Kuriyan, J. (1996) *Cell* **85**, 931-942
52. Lohmann, V., Korner, F., Koch, J., Herian, U., Theilmann, L., and Bartenschlager, R. (1999) *Science* **285**, 110-113
53. Gosert, R., Egger, D., Lohmann, V., Bartenschlager, R., Blum, H. E., Bienz, K., and Moradpour, D. (2003) *J.Virol.* **77**, 5487-5492
54. Dimitrova, M., Imbert, I., Kieny, M. P., and Schuster, C. (2003) *J Virol.* **77**, 5401-5414
55. Neddermann, P., Clementi, A., and De Francesco, R. (1999) *J.Virol.* **73**, 9984-9991
56. Asabe, S. I., Tanji, Y., Satoh, S., Kaneko, T., Kimura, K., and Shimotohno, K. (1997) *J.Virol.* **71**, 790-796
57. Ferguson, M. R., Fan, X., Mukherjee, M., Luo, J., Khan, R., Ferreon, J. C., Hilser, V. J., Shope, R. E., and Fox, R. O. (2004) *Protein Sci.* **13**, 626-632
58. Archer, S. J., Camac, D. M., Wu, Z. J., Farrow, N. A., Domaille, P. J., Wasserman, Z. R., Bukhtiyarova, M., Rizzo, C., Jagannathan, S., Mersinger, L. J., and Kettner, C. A. (2002) *Chem.Biol.* **9**, 79-92

FOOTNOTES

*We are grateful to Dr. Bruce Malcolm of the Schering Plough Research Institute for review of the manuscript, and for generously providing SCH6 and purified BK strain scNS3 proteases. We are also grateful to Dr. Wayne Bolen for helpful discussions, and Xiao Lian Liang and Dr. Bo Xu for assistance in protein expression and purification. Supported in part by grants from the National Institutes of Health: U19-AI40035 and R21-DA018054. K.L. is the John Mitchell Hemophilia of Georgia Liver Scholar of the American Liver Foundation. JCF is supported by The James W. McLaughlin Endowment. ACF is a recipient of a Sealy Center for Structural Biology Predoctoral Fellowship.

¹**Abbreviations:** HCV, Hepatitis C Virus; NS, Non-structural; GdnHCl, Guanidinium hydrochloride; IPTG, Isopropyl-D-thiogalactoside; SDS, Sodium dodecyl sulfate; PAGE, Polyacrylamide gel electrophoresis; OD, Optical density; DTT, Dithiothreitol; β -MeOH, β -mercaptoethanol; PDB, Protein Data Bank; HPLC, High performance liquid chromatography; FPLC, Fast performance liquid chromatography; TFA, Trifluoroacetic Acid; Sos, Son of Sevenless; SH3, Src-homology 3; PPII, Poly-proline II

Keywords: HCV; NS3 protease; TRIF; TLR; poly-proline helix; PPII; Sos; enzyme kinetics; viral mechanism

FIGURE LEGENDS

Figure 1. Proteolysis of TRIF by a full-length single-chain HCV NS3 protease-helicase. **(A).** SDS-PAGE analysis of the products of full-length TRIF cleavage by full-length scNS3 (FL scNS3) containing both the protease and helicase domains (lanes 1-4) and scNS3 (protease domain only, lanes 5-8). SCH6 is specific peptidomimetic inhibitor of the NS3 protease. **(B).** Three dimensional plot of the HPLC chromatogram (A_{220}) showing the decrease in the substrate TRIF p372 peptide (PSSTPC/SAHL) and increase of product peaks as a function of reaction time.

Figure 2. Comparison of the efficiency of cleavage of the full length TRIF versus the TRIF p372 peptide by scNS3. **(A).** SDS-PAGE analysis showing extent of full-length TRIF (4 μ M) cleavage at the indicated concentrations of scNS3. Lanes 7 and 8 are reactions mixes containing either scNS3 only (5 μ M) or TRIF only (4 μ M) respectively. **(B).** HPLC chromatogram showing the decrease in residual intact TRIF p372 peptide with increasing scNS3 concentrations (1, 2, 5, 8.3, 16.7 μ M). **(C).** EC_{50} determination using data from panel A (■) and B (○) fitted to the Hill equation (see Methods).

Figure 3. Presence of poly-proline II helices in TRIF. **(A).** Amino acid sequence of TRIF indicates an abundance of proline residues (bold face). The 17 amino acids surrounding the NS3/4A cleavage sequence, including the upstream poly-proline stretch, is underlined (\downarrow indicates cleavage site as determined previously (13)). **(B).** CD spectra of TRIF suggest a significant fraction of the protein adopts a poly-proline II helix (PPII) conformation.

Figure 4. The HCV NS5A protein contains a highly conserved Sos SH3-binding domain sequence (PPVPPRR). **(A).** At the top is shown an alignment of the murine Sos homolog 1 sequence, NS5A sequence from the genotype 1b BK virus (the source of NS3/4A protease used in this study) and genotype 1a H77c virus, and TRIF (GenBank accession numbers are shown) in the regions surrounding the putative SH3-binding domains. The SH3-binding domain motif is boxed, and consists of Φ PX Φ PX+ where Φ is a hydrophobic residue, X is variable, and + is generally basic (46). Below is

shown as schematic representation of the viral NS5A domain structure highlighting the NS4A interaction site (56) as well as the Sos homology sequence that has been shown to interact with the cellular Grb2 SH3 domain *in vivo*. **(B)**. Stereo view showing the characteristic poly-proline II (PPII) conformation of the Sos peptide (PPPVPPR; PDB entry 1SEM, chain C).

Figure 5. Interaction of the Sos peptide with the HCV NS3 protease, as monitored by fluorescence binding studies and HPLC protease inhibition kinetics. **(A)**. Binding of the Sos peptide to the HCV scNS3 protease (BK strain) as monitored by fluorescence spectroscopy. **(B)**. Inhibitory activity (IC_{50} determination) of the Sos peptide on the NS3-mediated proteolysis of the FRET peptide substrate (see Methods for details).

Figure 6. Interaction of the Sos peptide with the HCV NS3 protease, as monitored by NMR spectroscopy. **(A)** Representative peaks showing chemical shift changes from the unliganded (red) to the liganded state, (blue). S139, one of the active site residues, show significant chemical shift perturbation. R123, localized in β C2, was found to shift significantly in previous studies of peptide inhibitors based on the viral substrate (33,37), but show little if any change with Sos binding. **(B)** Plot of the magnitude of chemical shift perturbation for each residue in the protein. The active site residues, S139, H57, and D81 show significant chemical shift changes (labeled). Other residues that show significant shifts include K136 (3_{10} helix), and A157 (one of the residues in the substrate binding site).

Figure 7. Stereo view of a ribbon model of the HCV NS3 protease (PDB entry 1NS3). The residues affected by Sos binding are colored red. These are localized in the 3_{10} helix, the substrate binding site (β E2), and the catalytic triad residues (H57, D81, S139). Prior NMR studies (50) indicate that peptide inhibitors based on viral substrates form an antiparallel β -strand with hydrogen bond interactions involving certain residues in β E2, suggesting this is the substrate binding site. Other residues in β F2 and β C2 are also affected in NMR chemical shift studies of these peptide inhibitors (50;58).

Figure 8. Influence of the poly-proline sequence in TRIF on cleavage of a peptide substrate and interactions with the active site of the NS3 protease. **(A)** Cleavage of TRIF peptides that do or do not contain the poly-proline sequence by scNS3 protease. The experiment was carried out as described in Experimental Procedures except that the reaction mix was incubated at 25°C for 60 min prior to analysis by HPLC. The scNS3 concentration was 5 μ M. **(B)** Plot depicting differences in scNS3 binding by a short TRIF P-side peptide lacking the poly-proline sequence (Ac-PSSTPC-OH), versus a lengthier TRIF peptide that includes this stretch of upstream residues (Ac-PPPPPPPPSSTPC-OH). Binding was ascertained by changes in NS3 fluorescence.

Table 1. Peptide substrates used in HPLC cleavage assays.

	P ₈	P ₇	P ₆	P ₅	P ₄	P ₃	P ₂	P ₁	P' ₁	P' ₂	P' ₃	P' ₄	P' ₅	P' ₆	P' ₇
5A/5B			E	D	V	V	α ^a	C	/	S	M	S	Y		
4A/4B	E	F	D	E	M	E	E	C	/	A	S	H	L	P	Y
4B/5A			E	C	T	T	P	C	/	S	G	S	W	L	R
TRIF 372			P	S	S	T	P	C	/	S	A	H	L		

^aα, aminobutyric acid

Table 2: Kinetics of scNS3 Protease-mediated Cleavage of Viral and TRIF-derived Peptide Substrates

Substrate ^a	k_{cat} (min ⁻¹)	K_m (μM)	k_{cat}/K_m (M ⁻¹ s ⁻¹)
5A/5B	105 ± 6	120 ± 10	15000 ± 1500
4A/4B	16.8 ± 1.2	1100 ± 100	250 ± 30
4B/5A	12.9 ± 0.3	3900 ± 100	55 ± 2
TRIF p372	14.5 ± 1.3	10300 ± 1200	23 ± 3

^a see Table 1

Figure 1

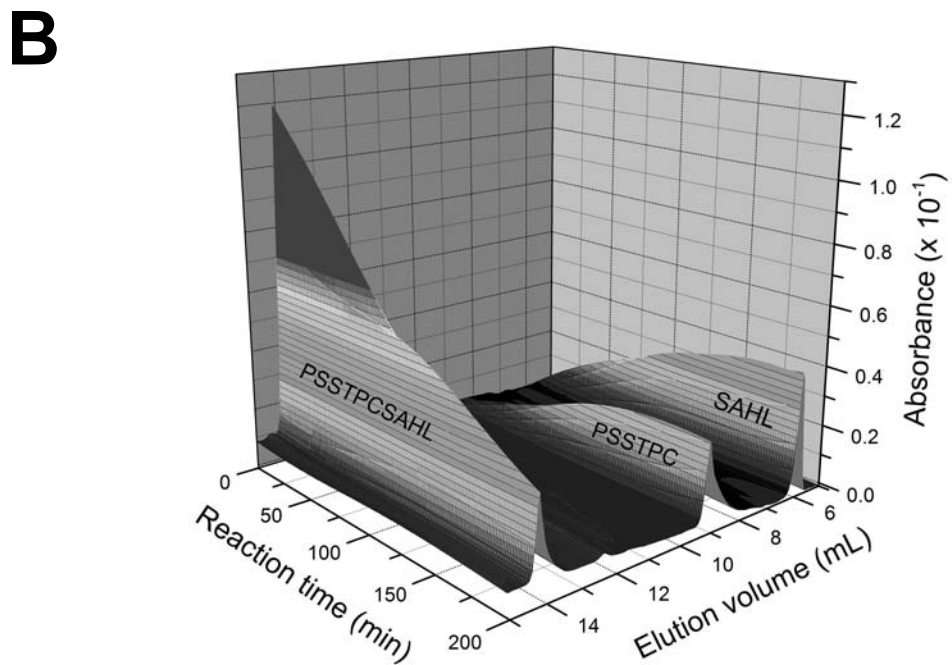
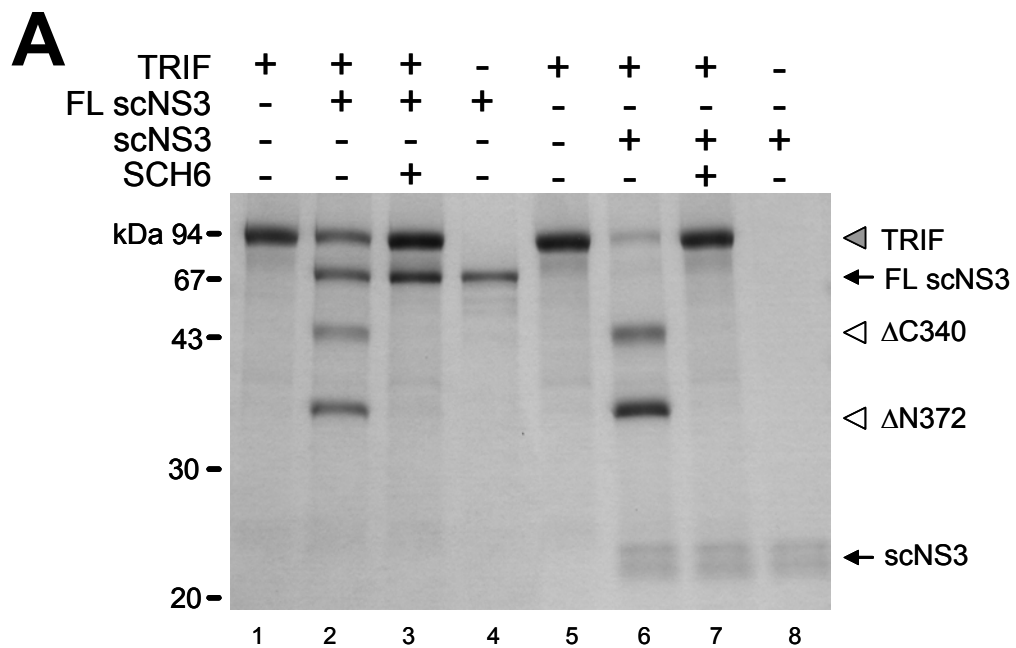


Figure 2

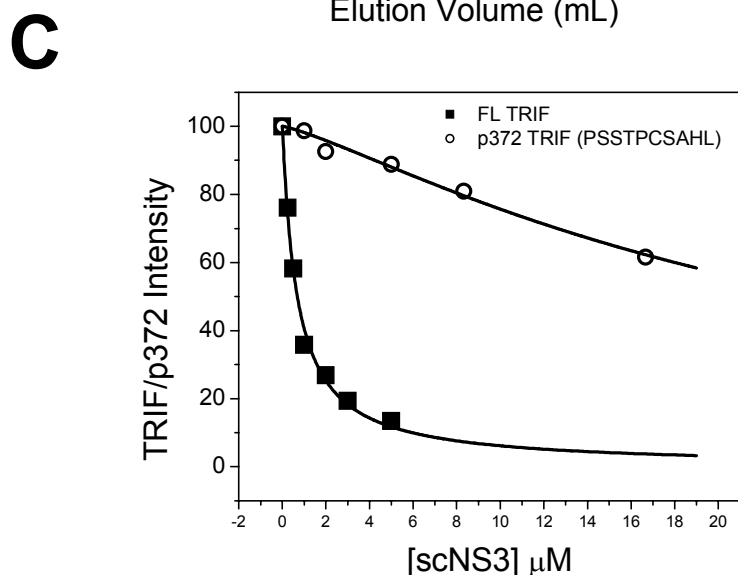
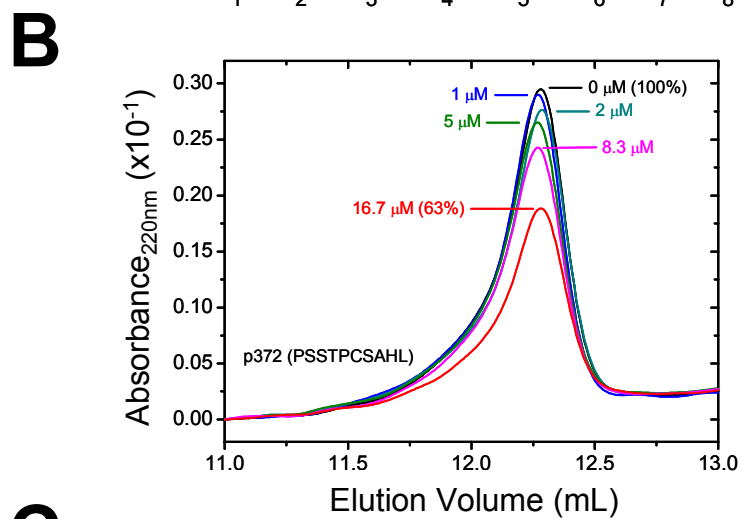
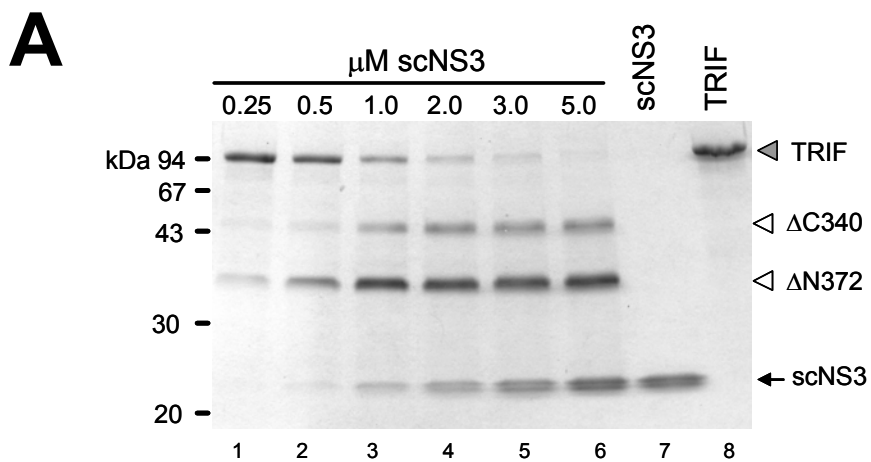


Figure 3

A

MACTG**PSLPS**AFDILGAAGQDKLLYLKHKLKT**PRPGC**QGDLLHAMVLLKLGQE
TEARISLEALKADAVARLVARQWAGVDSTED**PEEPPD**VSVAVARLYHLLAEEKL
C**PASLRD**VAYQEAVRTLSSRDDHRLGELQDEARNRCGWDIAGD**PGS**IRTLQSNL
GCL**PPSSALPSG**TRSL**PRPID**GVSDWSQGCSLRSTGS**PASLAS**NLEISQS**PTMP**
FLSLHRS**PHGPS**KLCD**DPQASLVPEPV**PGGCQE**PEEMSWPPS**GEIAS**PPELPSS**
PPPGLPEVAPDATSTGL**PDTPAAP**ETSTNY**PVECTEGS**AG**PQSLPLP**ILE**PVKN**
PCSVKDQTPLQLSVEDTTS**PNTKPCPPTPTT**PETS**PPPPPPPS**ST**PC**↓SAHLT
PSSLFPSSLESSSEQKFYNFVILHARADEHIALRVREKLEALGV**PDGAT**FCEDF
QV**PGRGELS**CLQDAIDHSAFIILLT**SNFDCRLSLHQVNQAMMS**NLTRQGS**PDC**
VI**PFLPLESSPA**QLSSDTASLLSGLVRLDEHSQIFARKVANTFK**PHRLQ**ARKAM
WRKEQDTRALREQS**QHLDGERMQAAALNAAYSAYLQSYLSYQAQMEQLQVAFGS**
HMSFGTGAPY**GARMPPGGQVPLGAPPPFPTWPGCQPPPL**HAWQAGT**PPPPSPQ**
PAAFPQSLPFPQSPA**FTAS****PAPPQS**PGLQ**PLI**IHHAQMVQLGLNNHMWN**QRGS**
Q**APEDKTQ**EAE

B

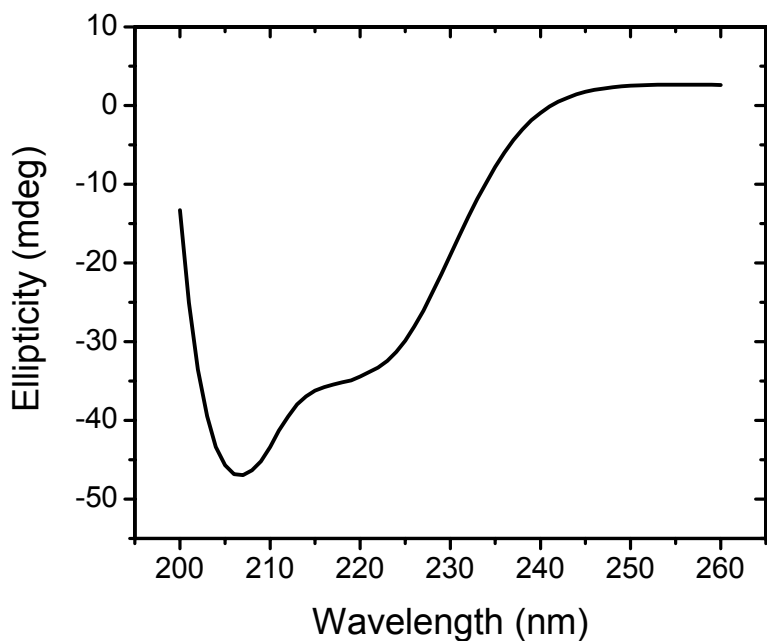
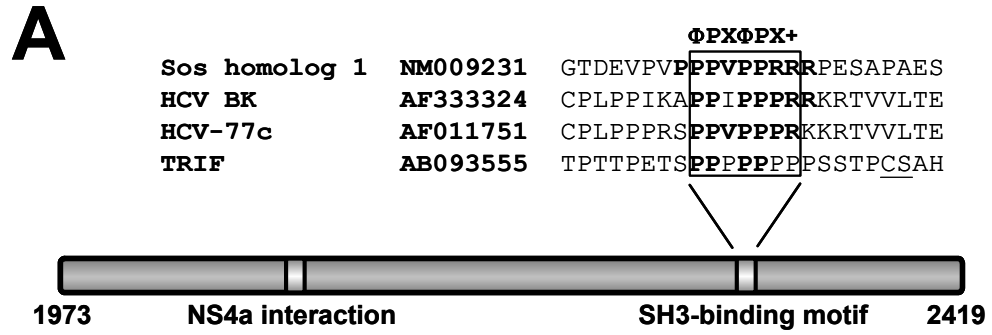


Figure 4.



B

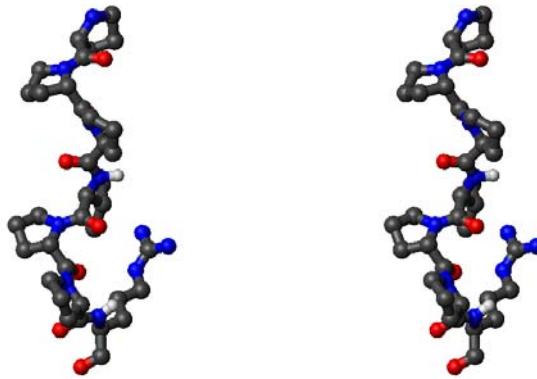


Figure 5.

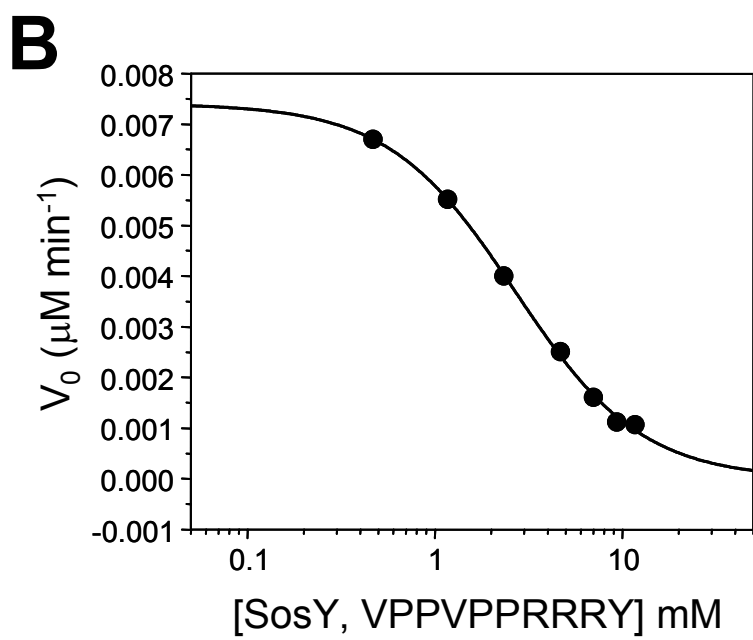
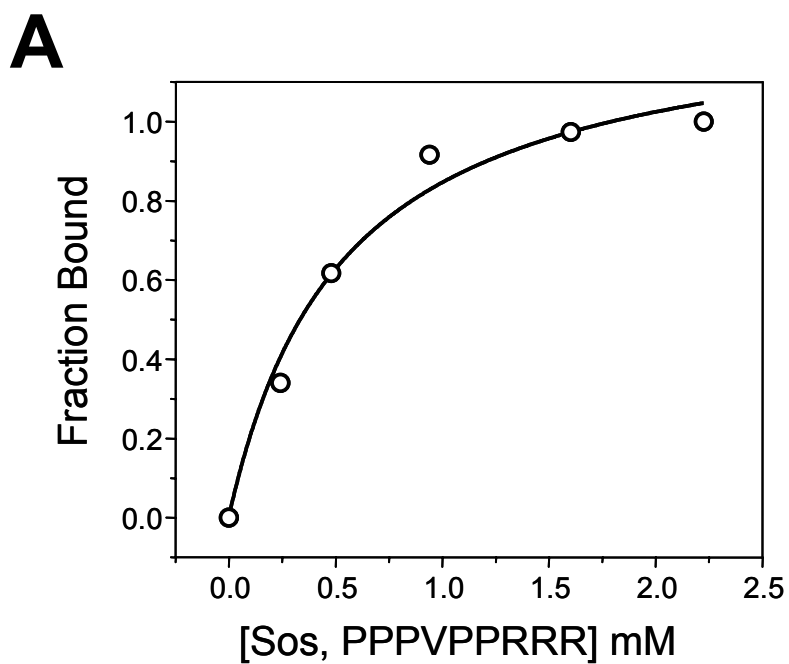
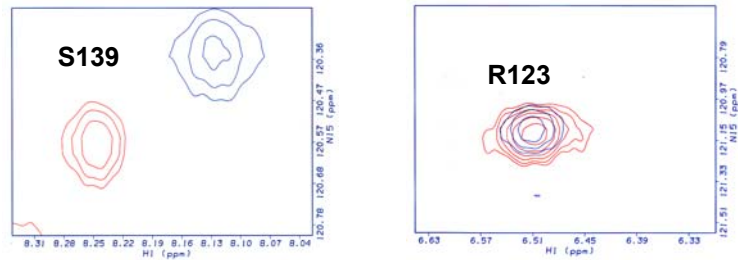


Figure 6.

A



B

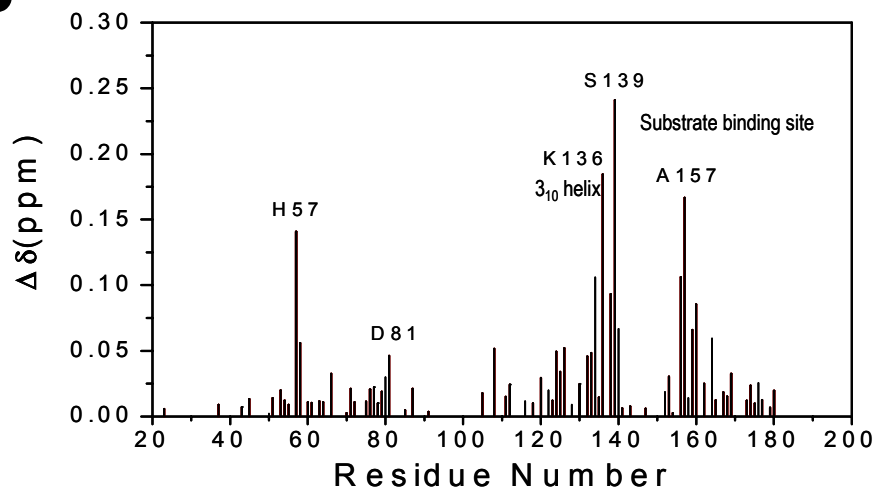


Figure 7.

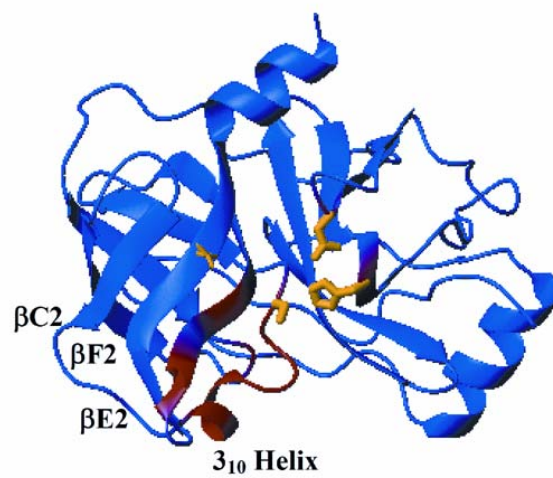
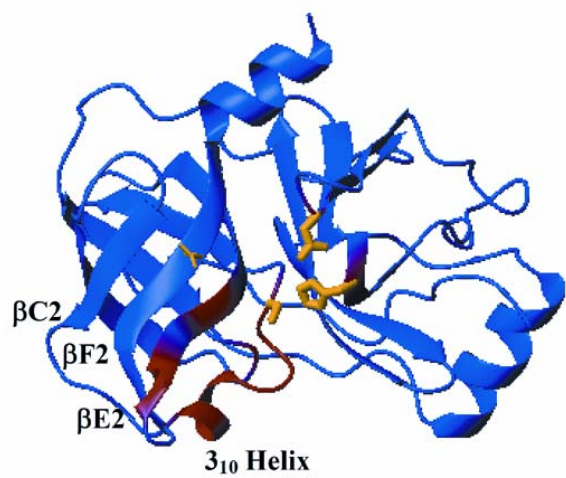


Figure 8.

