Plasmid constructs

For expression of NS3 protease (amino acid 1 to 180) an E. coli codon optimized synthetic nucleotide sequence harboring Nde I and Bam HI recognition sites at the 5’ and 3’ ends was used. The nucleotide sequence encodes for 7 histidine residues, a tobacco etch virus (TEV) protease cleavage site, the NS4A cofactor, a linker (GSGS) and NS3 protease (His7-NS4A-NS3) domain inserted into Nde I and Bam HI sites of pET11a vector. The N-terminal oligohistidine (His7)-TEV cleavage site fusion was incorporated to facilitate the purification of the protein and tag removal using TEV protease to produce the native protein for downstream applications.

Expression of full-length NS3 and NS3 protease in *E. coli*

For expression of full-length His6-NS3 (contain a serine protease and an RNA helicase) competent cells of BL21-CodonPlus (DE3)-RIL and E. coli BL21 (DE3) were transformed with pET11a-His6-NS3. For the expression of NS3 protease, pET11a-His7-NS4A-NS3 was transformed into only E. coli BL21 (DE3) cells considering the codon optimized nucleotide sequence of NS3 protease. After overnight growth, a single isolated colony was picked from the LB agar plate and transferred into 1.5 mL of LB medium supplemented with 125 µg/mL ampicillin and were grown at 37 °C overnight (~16 h) in an orbital shaker with 225 rpm shaking speed. Then, the cultures were diluted 100-fold (1% inoculum) in LB medium supplemented with 125 µg/mL of ampicillin and grown at 37 °C in a shaking incubator (225 rpm) until OD600 rises upto 0.5-0.6. The cultures were then incubated in cold room (at 4 °C) for 30 min and supplemented with 100 µM zinc chloride, and expression of full length NS3 and NS4A fused NS3 protease domain was induced by adding 1 mM isopropyl-β-D-thiogalactoside (IPTG) overnight (~16 h) at 18 °C and for 4 hours at 30 °C, respectively. Cells containing overexpressed proteins were harvested by centrifugation at 12000xg (Beckman Centrifuge) at 4 °C and the cells were stored at -80 °C. To check expression of proteins at small scale, bacterial protein extraction reagent (B-PER, ThermoFisher scientific) was used for cell lysis. Expression of target proteins was then analyzed on NuPAGE 4-12% Bis-Tris gel using the X-Cell Surelock Mini-cell electrophoresis system (Invitrogen).

Purification of full-length NS3 and NS4A-NS3 protease domain

All the protein purification procedures were carried at 4 °C using Bio-Rad NGC system. The cells lysis was done by using APV-1000 homogenizer (Invensys APV Products, Albertslund, Denmark) at ~10,000 psi, and centrifuged for 30 mins at 30,000 × g. The supernatants for both proteins were filtered separately with polyethersulfone membrane (0.45 μm) and then loaded onto His Trap column (GE healthcare) already equilibrated with their respective buffer A. Both the columns were washed to baseline with buffer A and eluted with a linear gradient of imidazole to 250 mM (His_6-NS3) and 500 mM (His_7-NS4A-NS3). Fractions containing recombinant His_6-NS3 or His_7-NS4A-NS3 proteins were pooled separately, concentrated separately using an Amicon Ultracel® 10 kDa cellulose membrane (EMD Millipore Corporation, Billerica, MA). High peak fractions of His_6-NS3 were dialyzed to decrease the imidazole concentration against dialysis buffer (25 mM pH 7.6 supplemented with 20% glycerol, 0.2 M NaCl, 10 mM β-mercaptoethanol and 0.4% Triton X-100) and the protein was concentrated using amicon membrane concentrator (cut-off 10kDa) upto 10 mL. To get the monodisperse oligomeric state, His_6-NS3 was further polished through gel filtration using an 120 mL HiPrep 16/60 Sephacryl S-200 column. On the otherhand, His_7-NS4A-NS3 was diluted using 25 mM HEPES (pH 7.5), 1 M NaCl, 10% glycerol buffer respectively, so that imidazole concentration reduced to about 25 mM and digested overnight at 4°C with His_6-tagged TEV protease. After 4°C digestion more His_6-tagged TEV protease was added to complete the cleavage at room temperature for 4 h. The digest was then transferred to a second His Trap column (GE healthcare) which was already equilibrated with buffer A and recombinant protein was obtained in the column effluent. After that, the effluent was incubated with 10 mM dithiothreitol for 16 hrs, concentrated using an Amicon Ultracel® 10 KDa cellulose membrane, and applied to a HiPrep 26/60 Sephacryl S-200 HR column (GE Healthcare Bio-Sciences Corporation, Piscataway, NJ) equilibrated in 20 mM HEPES (pH 7.5), 500 mM NaCl, 10% glycerol, 2mM TCEP (tris(2-carboxyethyl) phosphine buffer). The peak fractions containing His_6-NS3 or NS4A-NS3 protease were pooled and concentrated to ~ 1.5 mg ml^{-1} (measured at 280 nm using a molar extinction coefficient of 71500 M^{-1} cm^{-1} (His_6-NS3) 18700 M^{-1} cm^{-1} (NS4A-NS3 protease domain) derived using the Expasy ProtParam tool). Both proteins were used directly for the assay or flash-frozen in liquid nitrogen and stored at −80 °C. The final product was highly pure (>95%) as judged on sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS

PAGE). Molecular weight of the proteins was confirmed by electrospray ionization mass spectrometry.

Preparation of extracts

The extraction procedure was carried out using four solvents: n-hexane, ethyl acetate, ethanol and 70% ethanol in water. Each citrus plant powder (1 g) was added to 100 mL of each solvent and solution kept at 150 rpm shaking (~25 °C) in darkness for 24 h. The supernatant was filtered and the extracts were then concentrated in a rotary evaporator (Buchi R 210) under vacuum at 30 °C and 200 rpm and then stored under N_2 in airtight brown glass vials at −20 °C until used.

Inhibition of HCV NS4A-NS3 protease activity by plant extracts

A total reaction mixture of 300 µL using assay buffer: [50 mM HEPES pH 7.5; 10 mM DTT; 40% glycerol; 0.4% triton X-100]. Initial screening was performed by adding 3.33 mg extract solutions in DMSO (1 mL). The final concentration of DMSO was 3% in each reaction. 0.5 nM HCV NS4A-NS3 protease using multichannel pipette initiated the reaction, and release of 1 µM HCV Protease FRET Substrate (RET S1) was measured by Fluorescence Plate Reader (THE SPARK®, TECAN). For each assay plate a solvent (DMSO) control was performed by measuring activity of HCV NS4A-NS3 (without plant extract). The background fluorescence of each extract was measured. The calibration curve was plotted by spiking known concentrations of free RET S1 substrate into the same volume of reaction mixtures containing the plant extract, buffer, and enzyme. Hence, the slope of this curve was compared with that obtained from free RET S1 concentrations in the absence of plant extract and substrate. Initial velocities (rates) were determined, and then percentage (%) inhibition was estimated using the formula: 100 x [a-b/a] where "a" is the fluorescence of reaction mixture without inhibitor and "b" fluorescence of reaction mixture with inhibitor.

LCMS analysis of extracts

Samples were filtered and injected to MS with a flow rate of 10 µL min^{-1} through direct syringe pump. Scanning was done at positive and negative ion modes with mass range *m/z* 50-2,000. Source voltage and capillary voltage in positive ion mode were adjusted at 4.2 kV and 35kV, respectively. While in negative ion mode, these values were set at -4.5 kV and -30kV. In both scan modes, capillary temperature was set at 280 °C, sheath glass flow rate (N_2) was set at 25 L.min^{-1}, while auxiliary gas flow rate was set 5 L min^{-1}. Tandem mass spectrometry was done at both modes, sample analytes were fragmented by employing collision induced dissociation (CID) energy value at 25 (percentage of 5V) or otherwise stated. Full scan and fragmentation data were acquired and processed with Xcalibur software. Further data analysis was conducted using ChemBioDraw Ultra 14.0. The compounds were identified by matching their fragmentation finger prints with literature data and reference standards.

Supplementary Figures

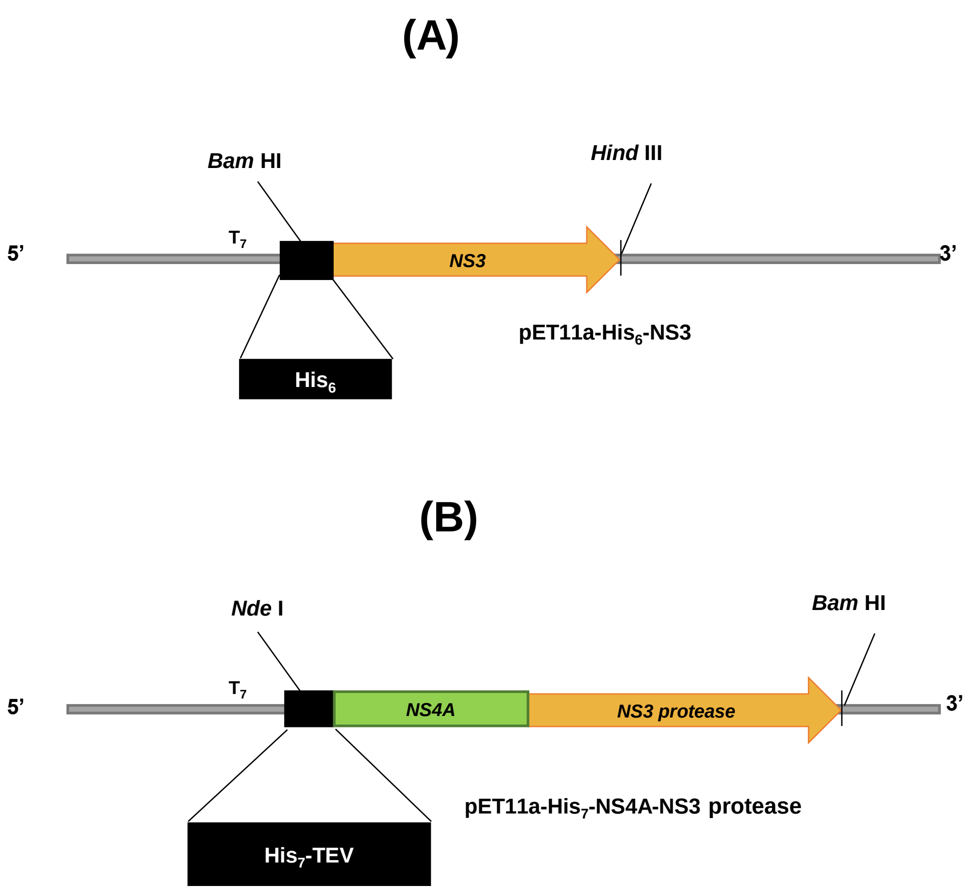


Figure S1: Strategy and graphical representation of pET11a-His$_6$-NS3 and pET11a-His$_7$-NS4A-NS3 expression vectors employed in this study, illustrating the position of encoded His-tag and TEV cleavages site sequences, key restriction sites, T$_7$ promoter and NS3 (helicase-protease) and NS4A (cofactor) - NS3 protease.

```
                         ---------His7---------..........................TEV...................------
  1     Nde I            H  H  H  H  H  H  H  E  N  L  Y  F  Q  S  V  G
  1 CATATGAAAAAACATCATCACCACCATCACCATGAAAATCTGTATTTCCAGAGCGTGGGT
  1             10        20        30        40        50
    ---------NS4A--------------------...........linker.........---------------
 21 S  V  V  I  V  G  H  I  E  L  G  G  S  G  S  G  I  T  A  Y
 61 AGCGTGGTGATTGTGGGCCACATTGAGCTCGGTGGCAGCGGTAGCGGCATTACCGCCTAC
 61             70        80        90        100       110
    ---------------------------NS3 protease---------------------
 41 A  Q  Q  T  R  G  L  F  G  T  I  I  T  S  L  T  G  R  D  K
121 GCCCAGCAGACCCGTGGTCTGTTTGGCACCATTATTACCAGCCTGACCGGCCGTGATAAA
121             130       140       150       160       170
    ---------------------------NS3 protease---------------------
 61 N  V  V  T  G  E  V  Q  V  L  S  T  A  T  Q  T  F  L  G  T
181 AACGTGGTGACCGGCGAAGTGCAGGTGCTGAGCACCGCGACCCAGACCTTTCTGGGCACC
181             190       200       210       220       230
    ---------------------------NS3 protease---------------------
 81 A  V  G  G  V  M  W  T  V  Y  H  G  A  G  S  R  A  L  A  S
241 GCGGTGGGCGGCGTGATGTGGACCGTGTATCATGGCGCGGGCAGCCGCGCGCTGGCGAGC
241             250       260       270       280       290
    ---------------------------NS3 protease---------------------
101 N  K  R  P  A  L  Q  M  Y  T  N  V  D  Q  D  L  V  G  W  P
301 AACAAACGCCCGGCGCTGCAGATGTATACCAATGTGGATCAGGATCTGGTGGGCTGGCCG
301             310       320       330       340       350
    ---------------------------NS3 protease---------------------
121 A  P  A  G  T  K  S  L  E  P  C  T  C  G  S  S  D  L  Y  L
361 GCGCCGGCGGGCACCAAAAGCCTGGAACCGTGCACCTGCGGCAGCAGCGATCTGTACCTG
361             370       380       390       400       410
    ---------------------------NS3 protease---------------------
141 V  T  R  E  A  D  V  L  P  A  R  R  R  G  D  S  T  A  S  L
421 GTGACCCGTGAAGCGGATGTGCTGCCGGCGCGTCGCCGTGGCGATAGCACCGCGAGCCTG
421             430       440       450       460       470
    ---------------------------NS3 protease---------------------
161 L  S  P  R  P  L  S  C  L  K  G  S  S  G  G  P  I  M  C  P
481 CTGAGCCCGCGTCCGCTGAGCTGCCTGAAAGGCAGCAGCGGCGGTCCGATTATGTGTCCG
481             490       500       510       520       530
    ---------------------------NS3 protease---------------------
181 A  G  H  V  V  G  I  F  R  A  A  V  C  T  R  G  V  A  K  A
541 GCGGGTCATGTGGTGGGTATTTTTCGTGCGGCGGTGTGCACCCGTGGCGTGGCGAAAGCG
541             550       560       570       580       590
    ---------------------------NS3 protease----------
201 L  Q  F  I  P  V  E  T  L  S  T  Q  A  R  S  *  *  Bam HI
601 CTGCAGTTTATTCCGGTGGAAACCCTGAGCACCCAAGCGCGTAGCTAATAAGGATCC
601             610       620       630       640       650
```

Figure S2: Synthetic nucleotide sequence (translated) showing *Nde* I / *Bam* HI recognition sequences, His7 tag, TEV cleavage site, NS4A (cofactor), linker (GSGS) and NS3 protease domain.

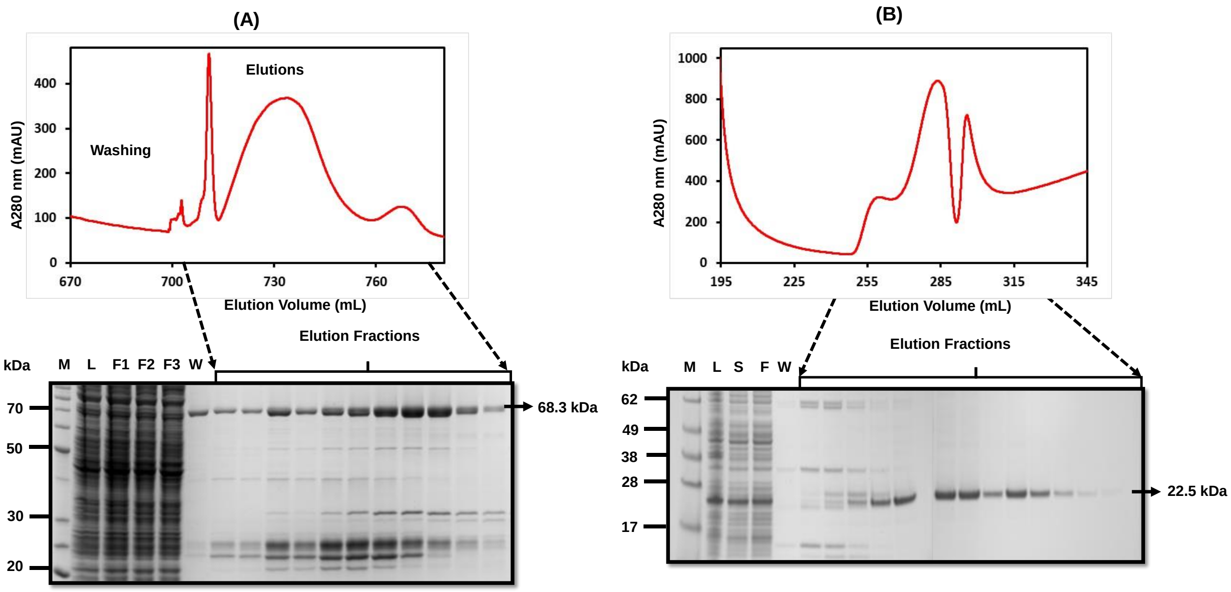


Figure S3. Purification of full-length His$_6$-NS3 and His$_7$-NS4A-NS3 protease domain by Nickel affinity chromatography. (A) Purification of the full-length His$_6$–NS3 protein from BL21-CodonPlus(DE3)-RIL/pET11a-His$_6$-NS3 cells. Lane M: Protein marker, BenchMark Unstained Protein Ladder (Invitrogen, Catalog #10747-012). Lane L: Cell lysate of the His$_6$–NS3 expressed cells (column load including soluble protein). Lane F1, F2 and F3 represent unbound fractions of lysate. Lane W: Column wash fraction. Lanes next to end: Elution fractions of purified His$_6$–NS3. **(B)** Purification of NS4A-NS3 protease domain from BL21(DE3)/pET11a-His$_7$-NS4A-NS3 cells. Lane M: Protein marker, Invitrogen Cat # 1891868 SeeBlue® Plus2 Prestained ladder. Lane L. Cell lysate of NS4A-NS3 protease domain expressed cells. Lane S: Soluble fraction of NS4A-NS3 protease domain expressed cells. Lane F: Unbound fractions of soluble lysate. Lane W: Column wash fraction. Lanes next to end: Elution fractions of purified NS4A-NS3 protease domain.

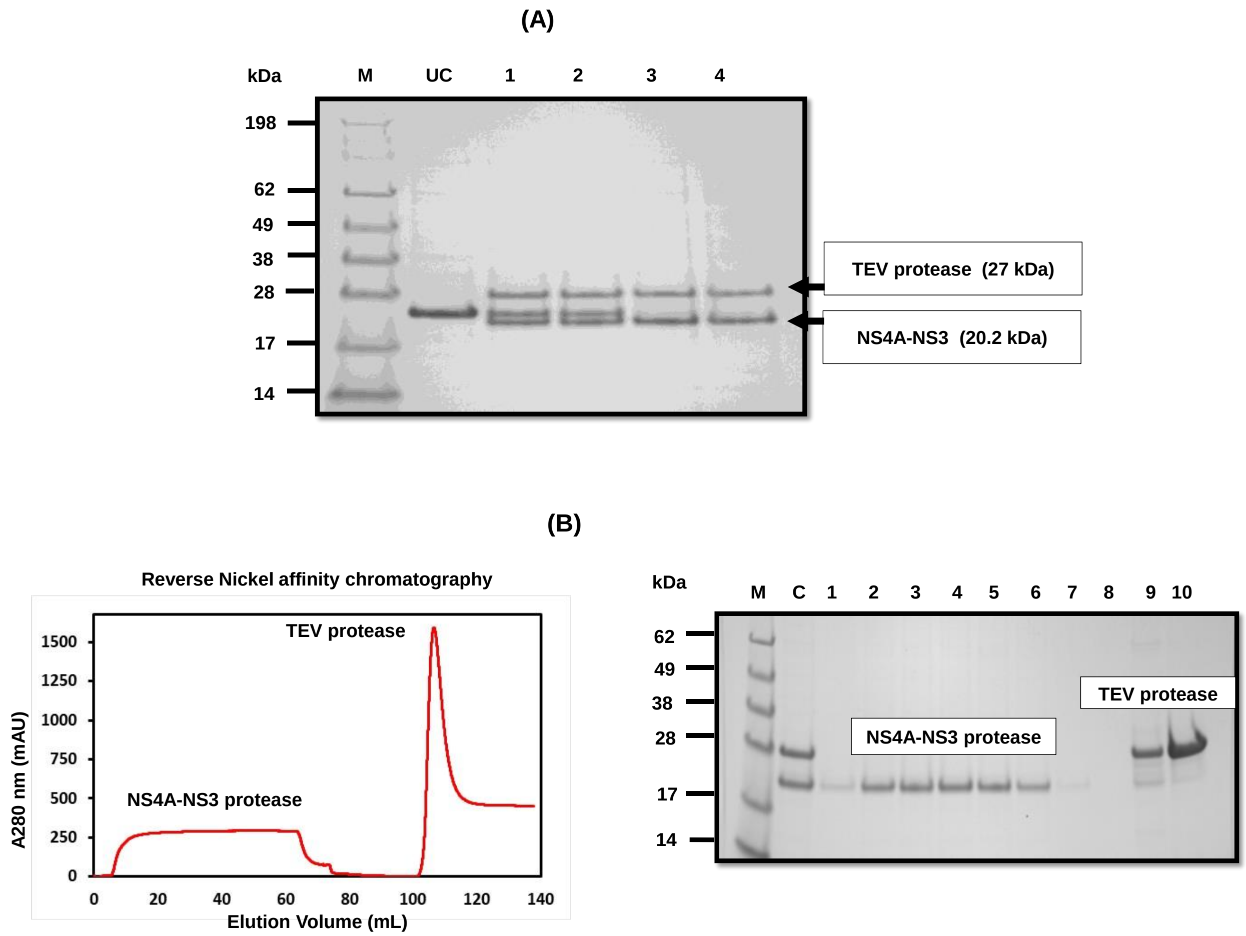
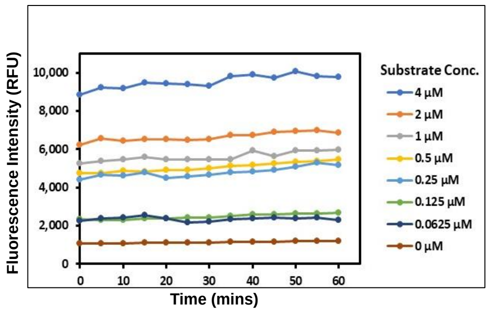
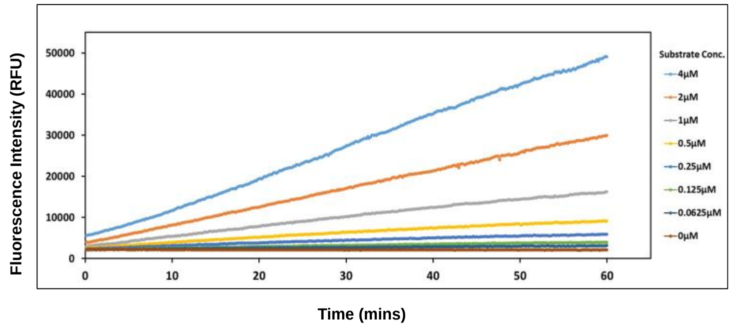


Figure S4: Production of the native NS4A-NS3 protease domain. (A) Analysis of cleavage of His$_7$-TEV-NS4A-NS3 protease domain with TEV protease. Lane M: Protein marker: Invitrogen Cat # 1891868 SeeBlue ® Plus2 Prestained ladder. Lane UC: Uncut His$_7$-TEV-NS4A-NS3 protease domain. Lane 1 and 2: Cleavage of His$_7$-TEV-NS4A-NS3 protease domain at 4 °C for 16 hours. Lane 3 and 4: Cleavage of His$_7$-TEV-NS4A-NS3 at room temperature for 4 hours. **(B)** Purification of the native NS4A-NS3 protease domain from cleavage mixture (His$_7$-TEV-NS4A-NS3 protease domain + TEV protease + His$_7$-TEV + NS4A-NS3 protease domain) by Nickel affinity chromatography.

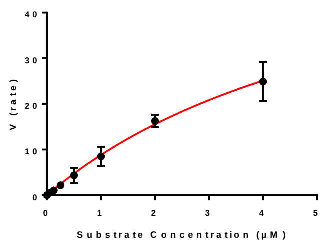
(A)



(B)



(C)



Kinetic values for NS4A-NS3 protease genotype 3a

K_m (µM)	k_{cat} (s^{-1})	k_{cat}/K_m ($µM^{-1}. s^{-1}$)
6.395 ± 2.049	0.096357688	0.01506766

Figure S5. of full-length His$_6$-NS3 and NS4A-NS3 protease domain. (A) Activity analysis of His$_6$–NS3 at different depsipeptide substrate concentration. **(B)** Activity analysis of NS4A–NS3 protease domain at different depsipeptide substrate concentration. **(C)** Concentration dependence of depsipeptide substrate cleavage by NS4A–NS3 protease domain and determination of kinetic parameters. Substrate cleavage reactions were performed as described in Methods section.