Supporting Information for:

Allosteric inhibition of the NS2B-NS3 protease from dengue virus

Muslum Yildiz[¶], Sumana Ghosh[¶#], Jeffrey A. Bell[§], Woody Sherman[§] and Jeanne A. Hardy[¶*]

[¶]Department of Chemistry, 104 LGRT, 710 N. Pleasant St., University of Massachusetts Amherst, MA 01003, USA

[#]Present Address: Vyome Biosciences, Plot No. 459 F.I.E, 1st Floor Patparganj Industrial Area Delhi, 110092 India

[§]Schrödinger, LLC, 120 West 45th Street, New York, New York 10036, United States

*corresponding author: phone (413) 545-3486; fax (413) 545-4490; hardy@chem.umass.edu

Running title: Dengue virus protease inhibition

METHODS

Dengue virus serotype 2 protease (DENV2 NS2B-NS3pro) expression constructs. The wild-type dengue virus serotype 2 protease (NS2B-NS3pro) construct encoding a synthetic gene (Celtek Bioscience) for *E. coli* codon-optimized N-terminally His_6-tagged dengue virus serotype 2 protease (DENV2 NS2B-NS3pro, Supplementary Fig. S8) gene was constructed by ligation into the XhoI/BamHI sites of the pET15b vector (Stratagene). This construct expresses a protein comprising amino acids His_6-NS2B (4395)-GGGGSGGGG-NS3pro (1185). Amino acid substitutions as described were introduced by the Quikchange® site-directed mutagenesis in this construct. All DNA sequences were validated by DNA sequencing (Genewiz).

Cysteine variant purification. NS2B-NS3pro expression constructs were transformed into the BL21 (DE3) T7 Express strain of *E. coli (*New England Biolabs). Cultures were grown in 2xYT media with ampicillin (100 mg/L, Sigma-Aldrich) at 37°C until they reached an optical density at 600 nm of 0.8. The temperature was reduced to 18°C and cells were induced with 1 mM IPTG (Anatrace) to express protein. Cells were harvested after 18 hrs expression and pelleted by 5 Krpm centrifugation in a SLC4000 rotor (Sorvall). Cell pellets were stored at –80°C, freeze-thawed and lysed by microfluidizer (Microfluidics, Inc.) in a buffer containing 50 mM Tris pH 8.5, 50 mM NaCl, 5% glycerol and 2 mM imidazole (Lysis buffer). Lysed cells were centrifuged at 17 krpm to remove cellular debris. The supernatant filtered in an 0.22-µm filter was loaded onto a 5-mL HiTrap Ni-affinity column (GE, Healthcare) equilibrated with lysis buffer. The column was washed with a buffer of 50 mM Tris pH 8.5, 50 mM NaCl, 5% glycerol and 10 mM imidazole, and the protein was eluted with the buffer containing 50 mM Tris pH 8.5, 50 mM NaCl, 5% glycerol and 100 mM imidazole.

The eluted fractions were diluted 6-fold with 50 mM Tris pH 8.5 buffer, to decrease the salt concentration. This protein sample was purified further by anion exchange chromatography. The partially purified sample was loaded onto a 5-mL HP High Q column (GE Healthcare) and

eluted with a linear NaCl gradient. NS2B-NS3pro eluted in a buffer of 50 mM Tris pH 8.5, 150 mM NaCl. The eluted protein was incubated with thrombin during dialysis against 50mM pH 8.5 Tris at 4°C overnight to cleave the His_6 tag. The following day the protein sample was loaded onto 5-mL HiTrap Ni-affinity column (GE Healthcare) to remove the cleaved His_6 tag and the thrombin. The His_6-free protein sample was concentrated to 5 mL and loaded to Hiload Superdex 75 26/60 size-exclusion column equilibrated with lysis buffer. Samples were eluted with lysis buffer and fractions were stored at –80°C in the lysis buffer conditions. The identity of the purified DENV2 NS2B-NS3pro and variants were analyzed by SDS-PAGE to be 99% pure and ESI-MS was used to confirm the mass of all proteins used.

Kinetics of cysteine-substituted variants. For kinetic measurements of NS2B-NS3pro activity, 300 nM-purified NS2B-NS3pro was monitored over the course of 20 min in a NS2B-NS3pro activity assay buffer containing 100 mM Tris pH 9.0. For substrate titrations, a range of 0-1mM fluorogenic substrate, GRR-AMC, (N-acetyl-Gly-Arg-Arg-AMC (7-amino-4-methylcoumarin); Ex 365 nm / Em 495 nm; International peptides) was added to start the reaction. Assays were performed in duplicate at 37°C in 100 µL volumes in 96-well microplate format using a Spectramax M5 spectrophotometric plate reader (Molecular Devices). Initial velocities versus substrate concentration were fit to rectangular hyperbola using GraphPad Prism (Graphpad Software) to determine kinetic parameters K_M and k_{cat}. Enzyme concentrations were determined by monitoring absorbance at 280 nm using a nano-drop (Thermo Scientific) instrument. The effect of probes on cysteine variants was determined by incubation of 1 µM of each NS2B-NS3pro variant with 100 µM of BAClMK, biarylamine or 10 µM of aldrithiol or DTNB for 1 hour prior to performing the activity assay as described above.

Mass spectrometry. NS2B-NS3pro A125C in Tris pH 8.5, 50mM NaCl and 5% glycerol was diluted in 0.1% formic acid to a final concentration of 5 μM in a 200 μl volume and analyzed on Qstar-LC electrospray ion trap mass spectrometer (Applied Biosystems) with an ESI source and positive ion polarity. A PROTO 300 C4 5 μM column (Higgins Analytical, Inc.) was used to desalt the protein on the LC prior to ionization in the mass spectrometer. The MS instrument system was equipped with an HP1100 HPLC system (Hewlett-Packard). Scanning was carried out between 600-1400 m/z. A mass range between 20 kDa and 30 kDa was deconvoluted to obtain corresponding sample mass. The expected mass for the unliganded form of A125C was 26329 Da; we observed 26332 Da as expected.

To validate the compound binding, 10 μM of NS2B-NS3pro A125C was incubated with 100 μM DTNB (5,5'-dithiobis-(2-nitrobenzoic acid)) or 100 μM BAClMK in Tris pH 8.5, 50 mM NaCl and 5% glycerol buffer for 1 hr to ensure full binding. After a 1 hr incubation, the sample buffer was exchanged for 0.1 % formic acid and analyzed as above. The expected mass for DTNB bound form of A125C was 26,527 Da; we observed the major species (>95%) as 26,526 Da, which clearly indicates covalent compound binding. Likewise, the expected mass for BAClMK-bound of A125C was 26,682 Da; we observed the major species (>95%) as 26,684 Da, which clearly indicates covalent compound binding.

Determination of molar binding ratio of DTNB and BAClMK

The total protein concentrations of NS2B-NS3pro WT, E19C, L31C, E86C and N105C were determined by measuring 280 nm absorbance and a molar extinction coefficient of 41970 M^{-1} cm^{-1}. Samples in 0.1 M pH 8.5 Tris, 1 mM EDTA were then incubated in 50 μM DTNB for 30 minutes. The concentration of sulfhydryl for each sample was determined by measuring the absorbance (412 nm) of the released yellow product of the DNTB reaction with free thiol, 2-nitro-5-thiobenzoate. The released 2-nitro-5-thiobenzoate concentration was calculated using

Beer's Law and the molar extinction coefficient for 2-nitro-5-thiobenzoate of 14,150 $M^{-1}cm^{-1}$. The concentration of total protein was compared to the concentration of released 2-nitro-5-thiobenzoate, to calculate the DTNB molar binding ratio.

To determine the molar binding ratio of BAClMK, NS2B-NS3pro variants in 0.1 M pH 8.5 Tris were incubated in 500 μM BAClMK for 1 hour. Afterward unreacted BAClMK was removed and samples were exchanged to a buffer that contained 1mM EDTA and 0.1 M pH 8.5 Tris using a NAP25 column (GE Healthcare). Each sample was then incubated in 50 μM of DTNB for 30 minutes to assess the presence of free thiol. Released 2-nitro-5-thiobenzoate was monitored at 412 nm and concentration was calculated as described above. The molar binding ratio of BAClMK was calculated to as the total protein minus the free thiols measured after BAClMK incubation.

Circular Dichroism (CD). Circular dichroism experiments were carried out on a J720 circular dichrometer equipped with peltier temperature controller (JASCO) at 20°C. The secondary structure contents of 5 μM WT NS2B-NS3pro, apo A125C and liganded A125C in 10 mM pH 8.5-phosphate buffer were observed by monitoring circular dichroism signal between 250 and 190 nm. A125C was labeled in lysis buffer (50mM pH 8.5 tris, 5% glycerol and 50 mM NaCl) with 50 μM DTNB (5,5'-dithiobis-(2-nitrobenzoic acid)) (Thermo Scientific) and buffer exchanged to 10 mM pH 8.5 phosphate buffer prior to the measurement.

Size Exclusion Chromatography. The oligomeric state of the NS2B-NS3pro WT or A125C unliganded or in various compound-bound states was determined by analysis on a Superdex 200 10/300 GL (GE Healthcare) gel-filtration column. NS2B-NS3pro samples were concentrated in 50 mM NaCl, 5% glycerol and pH 8.5 Tris in a 3K molecular weight cutoff membrane concentrator (Sartorium Stedim Biotech) to 2 mg/mL. Samples were filtered and loaded onto the column. Protein was eluted with 50 mM NaCl, 5% glycerol and pH 8.5 Tris buffer. Peak

fractions were identified by monitoring absorbance at 280 nm and the eluted proteins were analyzed by SDS-PAGE. Molecular weight standards from the gel-filtration calibration kit LMW (GE Healthcare) were run prior to experiment with the same conditions. A standard curve was drawn to determine the oligomeric state of NS2B-NS3pro. According to standard curve the expected retention volume for a monomer of NS2B-NS3pro (26 kDa) was 15.40 mL.

Dynamic light scattering (DLS). 2 mg/mL of unliganded or DTNB-labeled A125C in 0.1 M Bis-tris propane pH 6.0 and 0.02 % w/v n-Octyl-b-D-glucoside at 20°C was analyzed on a Zetasizer Nano ZS (Malvern) with a temperature control and a 633 nm He-Ne laser for backscattering at 173°. 15 measurements of 10s duration were averaged per analysis. The distributions of the mean apparent translational diffusion coefficients (D_T) were calculated by fitting the DLS autocorrelation functions using non-negative constrained least squares (NNLS). The distribution of apparent sizes d_h was obtained from the distribution of mean apparent translational diffusion coefficient (D_T) where k is the Boltzmann constant and η is the solvent viscosity, which was assumed to be that of water.

$$R_{app} = kT / (6 \pi \eta D_T)$$

Intrinsic Tryptophan Fluorescence. A125C was incubated with 10 μM DTNB in Tris pH 8.5, 50 mM NaCl and 5% glycerol buffer for 1 hr. The sample was buffer was exchanged to lysis buffer (50mM pH 8.5 Tris, 50mM NaCl and 5% glycerol) using a NAP5 column (GE Healthcare). The intrinsic tryptophan fluorescence of 500 nM WT or A125C, unliganded or in complex with DTNB, in lysis buffer was measured (Ex. 290/Em. 300-400 nm) on a Quantum Master 30 spectrofluorometer (PTI) as a function of temperature ranging from 25 to 85 °C. Temperature was controlled via circulating water bath. Resultant data were fit to a Gaussian hyperbole and the wavelength at the maximum absorption (λ_{max}) found by calculating the point

at which the 1st derivative was equal to 0 (Mathematica, Wolfram Research, Inc). λ_{max} shifted from 338 nm (100% folded) to 350 nm (0% folded) as temperature increased from 25 to 85°C . The fraction of folded NS2B-NS3pro was calculated at each temperature by monitoring the λ_{max} value.

Crystallization and Data Collection. Purified NS2B-NS3pro, wild-type or A125C, in 50 mM Tris pH 8.5, 5% glycerol, and 50 mM NaCl was concentrated using Amicon Ultrafree 3K NMWL membrane concentrators (Millipore) to 15 mg/mL. The crystals grew in 0.1 M Sodium Acetate pH 5.5, 40% PEG 200 in 1 week at 4°C temperature in a hanging drops. For the pH 8.5 WT or A125C structures, crystals were manually transferred to new mother liquor containing Tris pH 8.5 and 40% PEG 200 and soaked for 12 hrs. For A125C pH 8.5 + DTNB or + BAClMK structures, A125C crystals equilibrated at pH 8.5 were soaked for 6 hrs in the new mother liquor containing 5 mM DTNB (5,5'-dithiobis-(2-nitrobenzoic acid)) (Thermo Scientific) or BAClMK. All crystals were flash frozen in liquid nitrogen (LN_2) and stored in LN_2. Complete data sets for all unliganded proteins were collected at X6A beamline at 1 Å Brookhaven National Laboratories National Synchrotron Light Source (Upton, NY). Data for A125C pH 8.5 + DTNB were collected using synchrotron radiation of 1.0Å at APS 24-ID-C. Indexing and integration were carried out with MAR XDS(*1*). Data were scaled using SCALA.(*2*)

Structure Determination. All data were processed and scaled in HKL2000 in the orthorhombic space group $C222_1$ with CCP4(*3*). The search model successfully used for molecular replacement using Phaser(*4*) was unliganded DENV2 NS2B-NS3pro (PDB ID 2FOM, 1.5 Å resolution). The initial model building in Coot(*5*) and NCS restrained refinement using CCP4i. The structures were refined using automated refinement in Refmac(*6*) and TLS to a final R/R_{free} as listed for the various structures in Table S1. For A125CpH 8.5 + DTNB we performed additional refinement using the program PrimeX,(*7*) which provides both the advantages of all-atom model refinement(*8*) as well as a specialized facility for investigating ligand binding. The

water model from the original REFMAC-refined structure, as well as approximately 10% of the residues with the highest real-space R-factors, were removed from the working model. Reciprocal-space minimization of coordinates and temperature factors, and hydrogen-bond optimization, were applied periodically as atoms were added to the model. Missing internal residues were built using the PrimeX automated loop refinement tool. Similarly, N-terminal and C-terminal ends of the model chains were extended as far as the electron density permitted. At this stage additional analysis was performed of the residual electron density, during which no feature consistent with DTNB binding was detected. Waters were added by PrimeX at the 4σ level on F_o-F_c maps, and were screened manually to include only those with spherical electron density and a hydrogen-bonding interaction with the protein. The additional refinement decreased the free R-factor by 2.9%.

Synthetic materials and methods. All solvents and reagents used were of highest purity available. 4-phenoxyaniline, 3-mercaptopropanoic acid, 3,3'-Dithiodipropionic acid were purchased from Acros Organics. ^{1}H-NMR spectra were recorded on Bruker 400 spectrometer and chemical shifts (δ) are reported in ppm downfield from the internal standard (TMS).

Synthesis of Biarylchloromethylketone (BAClMK, 1, Fig S2A**)**

Synthesis of *N*-(4-Phenoxyphenyl)succinamic acid ethyl ester (**3**, Supplementary Fig. S2A). 4-phenoxyaniline (**2** (Supplementary Fig. S2A), 0.5g, 2.7 mmol) was dissolved in 3 mL dichloromethane (DCM). The solution was cooled and triethylamine (750 μl, 5.4 mmol) was added followed by dropwise addition of ethylsuccinylchloride (400 μl, 2.8 mmol) while maintaining the temperature of the reaction mixture at 0°C. The resultant solution was stirred at 0°C for 1-2 hr and at room temperature for 6-7 h. After complete consumption of the starting material, the reaction mixture was diluted with DCM and the product was successively washed with 0.5 N HCl, water and finally with brine solution. The organic phase was dried over sodium

sulfate (Na_2SO_4) and evaporated under vacuo to obtain **3** (Supplementary Fig. S2A) as a brown solid (0.507 g, 60%) with 90-95% purity. ^{1}H-NMR ($CDCl_3$, 400 MHz) δ (ppm) 1.25 (t, J = 7.2 Hz, 3H), 2.64 (t, J = 6 Hz, 2H), 2.74 (t, J = 6.5 Hz, 2H), 4.15 (q, J_1 = 7.1 Hz, 2H), 6.96 (d, J = 8.7 Hz, 4H), 7.06 (t, J = 7.4 Hz, 1H), 7.3 (t, J = 8.2 Hz, 2 H), 7.45 (d, J = 8.8 Hz, 2H), 7.68 (1H, NH).

Synthesis of *N*-(4-Phenoxyphenyl)succinamic acid (**4,** Supplementary Fig. S2A). Compound **3** (Supplementary Fig. S2A) (0.5 g, 1.6 mmol) was dissolved in 10 mL THF; and 20% sodium hydroxide solution (NaOH) (5 mL) was added into the reaction mixture. The resulting solution was refluxed for 3 hr and cooled to room temperature. The solvent was removed under vacuum and the resulting aqueous solution was acidified with 2 N HCl. The precipitate formed upon acidification was filtered, washed with minimum water and dried to obtain 0.4 g of (90%) brown solid. ^{1}H-NMR (DMSO (d_6), 400 MHz) δ (ppm) 2.57 (t, J = 6.5 Hz, 2H), 2.68 (t, J = 6.4 Hz, 2H), 6.9 (t, J = 8.7 Hz, 4H), 7.01 (t, J = 7.4 Hz, 1H), 7.25 (t, J = 8.4 Hz, 2 H), 7.45 (d, J = 8.9 Hz, 2H), 7.68 (1H, NH). ^{1}H-NMR (CH_3OH (d_4), 400 MHz) δ (ppm) 2.6 (s, 4H), 6.9 (t, J = 8.9 Hz, 4H), 7.01 (t, J = 7.4 Hz, 1H), 7.25 (t, J = 8.4 Hz, 2 H), 7.45 (d, J = 8.9 Hz, 2H), 7.68 (1H, NH).

Synthesis of *N*α-((*S*)-4-chloro-3-oxobutan-2-yl)-*N*δ-(4-phenoxyphenyl)succinamide (**1,** Supplementary Fig. S2A). Compound **4** (Supplementary Fig. S2A) (0.12 g, 0.42 mmol) was dissolved in THF (5 mL) and allowed to react with isopropylchloroformate (66 µl, 0.5 mmol) in the presence of *N*-methylmorpholine (60 µl, 0.55 mmol) for 1.5 hr at -20°C to activate the acid functionality. The amine salt of the L-alanine chloromethylketone (**5**, 0.094 g, 0.46 mmol) in DMF solution was added into the resulting solution by maintaining the temperature at -20 °C. The reaction mixture was stirred for another 1 hr at -20°C and allowed to stir at room temperature for overnight. The solvent was evaporated and extracted with ethylacetate. The ethylacetate fraction was washed successively with saturated sodium bicarbonate ($NaHCO_3$), 0.5

N HCl and finally with brine solution. The crude product obtained as brown solid was then purified by flash column chromatography using 4% MeOH/CH_2Cl_2 as the eluent to obtain **1** as a light brown solid (0.05 g, 31 with 98-99% purity. ^{1}H-NMR ($CDCl_3$ + CD_3OD, 400 MHz) δ (ppm) 1.28 (d, J = 6.5 Hz, 3H), 2.52 (d, J = 6.0 Hz, 2H), 2.56 (d, J = 5.7 Hz, 2H), 4.23 (dd, J_1 = 12.7 Hz, J_2 = 16.04 Hz, 2H), 4.5 (q, J_1 = J_2 = 7.2 Hz, 1H), 6.86 (d, J = 8.0 Hz, 4H), 6.98 (t, J = 7.4 Hz, 1H), 7.2 (t, J_1 = J_2 = 7.9 Hz, 2H), 7.4 (d, J = 8.9 Hz, 2H), 7.7 (1H, NH)., ^{13}C-NMR ($CDCl_3$ + CD_3OD), 120 MHz) δ (ppm) 14.56, 29.4, 30.52, 45.62, 51.8, 117.35, 118.4, 120.8, 122.1, 128.74, 133.2, 152.5, 156.8, 170.45, 172.55, 201.31., MS (FAB) 389.1(M+H), 388.8 (calculated mass).

Synthesis of Biarylthiol (1, Fig S2B**)**

Synthesis of 3-(tritylthio)propanoic acid, **3** (Supplementary Fig. S2B). In an air-dried flask (3.0 mL, 0.034 mole) 3-mercaptopropanoic acid was dissolved in 60 mL DCM. Trityl chloride (9.53 g, 0.034 mole) dissolved in THF, was slowly added into the reaction mixture at RT. After 15 min a white solid started forming and the resultant reaction mixture was allowed to stir for 14 hr at room temperature. Finally, the reaction mixture was filtered and washed three times with DCM. The filtrate was concentrated, which led to the formation of a solid. The solid was washed with minimal amount of dichloromethane. The combined solid was almost 95% pure and obtained as 90% isolated yield. The obtained product was directly used for the next step without further purification. ^{1}H-NMR ($CDCl_3$, 400 MHz) δ (ppm) 2.24 (t, J = 7.3 Hz, 2H), 2.46 (t, J = 7.3 Hz, 2H), 7.21-7.23 (m, 2H), 7.26-7.3 (m, 8H), 7.41-7.43 (m, 5H).

Synthesis of *N*-(4-phenoxyphenyl)-3-(tritylthio)propanamide, **4** (Supplementary Fig. S2B). 3-(tritylthio)propanoic acid (0.15 g, 0.92 mmol) was dissolved in 5 mL THF. The reaction mixture was cooled to -15°C to -20°C and a mixture of *N*-methylmorpholine (0.26 mL, 2.4 mmol) and isopropyl chloroformate (0.14 mL, 1.1 mmol) were added dropwise while maintaining the

temperature at -20°C. The resultant solution was further stirred at -20°C for 90 min to complete acid activation step. 4-phenoxybenzamine (0.17 g, 0.92 mmol) was dissolved in anhydrous THF and was added into the reaction mixture at -20°C. The resultant mixture was stirred at -20°C for 1 hr and slowly allowed to reach at room temperature. The solvent was evaporated and extracted with ethylacetate. The ethylacetate fraction was washed successively with saturated $NaHCO_3$, and finally with brine solution. The crude product was repeatedly washed with a minimal amount of ethylacetate and the filtrate was removed. By repeating this process three to four times a 95-98% pure brown solid product was obtained as (0.33 g) 70% isolated yield. ^{1}H-NMR ($CDCl_3$, 400 MHz) δ (ppm) 2.14 (t, *J* = 7.2 Hz, 2H), 2.6 (t, *J* = 7.2 Hz, 2H), 6.94-6.97 (m, 4H), 7.07 (t, *J* = 7.4 Hz, 1H), 7.21-7.31 (m, 12H), 7.4 (d, *J* = 7.4 Hz, 2H), 7.45 (d, *J* = 7.6 Hz, 5H).

Synthesis of 3-Mercapto-*N*-(4-phenoxyphenyl)propanamide, **1** (Supplementary Fig. S2B). In an air-dried flask, 3.1 mL trifluoroacetic acid was added to a stirring solution of **4** (0.096g, 0.186 mmol) in 6.2 mL DCM at room temperature. The yellow color produced was then quenched by the addition of 60 μl of Triethylsilane (TES). The reaction mixture was stirred at room temperature for 2 hr and then concentrated. After removal of the volatiles under vacuum, the solid residue was washed successively with hexane to remove triphenylmethane. Then the solid was dissolved in minimal volume of ethylacetate and the product was precipitated by using an excess of diethylether. This process was repeated two to three times to afford a white-colored solid at a 70% (0.07 g) isolated yield. ^{1}H-NMR (CD_3COCD_3, 400 MHz) δ (ppm) 1.5 (s, 1H), 2.7 (t, *J* = 7.2 Hz, 2H), 2.97 (t, *J* = 7.0 Hz, 2H), 6.93-6.97 (m, 4H), 7.08 (t, J_1 = 7.9 Hz, J_1 = 7.0 Hz, 1H), 7.32-7.36 (m, 2H), 7.67-7.69 (d, *J* = 7.9 Hz, 2H), 9.4(s, 1H).

References:

1. Kabsch, W. (2010) Xds, *Acta Crystallogr D Biol Crystallogr 66*, 125-132.
2. Collaborative Computational Project, N. (1994) The CCP4 suite: programs for protein crystallography, *Acta Crystallogr D Biol Crystallogr 50*, 760-763.
3. (1994) The CCP4 suite: programs for protein crystallography, *Acta Crystallogr D Biol Crystallogr 50*, 760-763.
4. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software, *J Appl Crystallogr 40*, 658-674.
5. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics, *Acta Crystallogr D Biol Crystallogr 60*, 2126-2132.
6. Murshudov, G. N., Skubak, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F., and Vagin, A. A. REFMAC5 for the refinement of macromolecular crystal structures, *Acta Crystallogr D Biol Crystallogr 67*, 355-367.
7. Bell, J., Cao, Y., Gunn, J., Day, T., Gallicchio, E., Zhou, Z., Levy, R., and Farid, R. (2012) PrimeX and the Schrödinger Computational Chemistry Suite of Programs, *International Tables for Crystallography Volume F: Crystallography of biological macromolecules F(18)*, 534-538.
8. Bell, J. A., Ho, K. L., and Farid, R. (2012) Significant reduction in errors associated with nonbonded contacts in protein crystal structures: automated all-atom refinement with PrimeX, *Acta Crystallogr D Biol Crystallogr 68*, 935-952.

Table S1. Crystallographic data collection and refinement statistics for NS2B-NS3pro variants

	WT pH 5.5	WT pH 8.5	A125C pH 5.5	A125C pH 8.5	A125C pH 8.5 + DTNB	A125C pH 8.5 + BAClMK
Wavelength (Å)	1.0	1.0	1.0	1.0	1.0	1.0
Diffraction Resolution (Å)[a]	50-1.46 (1.49-1.46)	115-1.53 (1.56-1.53)	50-2.37 (2.41-2.37)	58-2.65 (2.7-2.65)	44-1.70 (1.73-1.70)	No diffraction
Measured Reflections (n)	180652	221559	31758	43218	60921	
Unique Reflections	36703	32669	9008	6357	18461	
Completeness (%)	99.4 (99.9)	99.3 (84.9)	96.6 (98.8)	99.2(98.4)	85.3 (79.2)	
Redundancy[a]	4.9 (4.3)	3.6 (2.3)	3.5 (3.2)	6.8 (6.6)	3.3 (4.6)	
$I/\sigma(I)$ [a]	36.0 (2.1)	4.6 (2.6)	8.4 (2.0)	9.3 (1.5)	30.9 (1.9)	
R_{sym}	0.065 (0.51)	0.086 (0.51)	0.145 (0.54)	0.17 (0.96)	0.11 (0.66)	
Space group	$C222_1$	$C222_1$	$C222_1$	$C222_1$	$C222_1$	
a (Å)	59.5	59.9	60.7	60.2	61.4	
b (Å)	62.0	62.5	62.1	61.0	62.1	
c (Å)	114.1	114.6	115.0	114.0	114.4	
$\alpha = \gamma = \beta$ (°)	90	90	90	90	90	
Refinement statistics						
No. waters	103	169	14	36	30	
R_{work}/R_{free} (%)	20.3/22.3	19.6/23.3	21.5/27	22.7/24.0	21.4/24.1	
RMSD bond length (Å)	0.028	0.011	0.015	0.013	0.013	
RMSD bond angle (°)	2.7	1.5	1.83	1.71	1.99	
Average B-factor ($Å^2$)	24.2	26.51	53.4	48.5	27.4	
Ramachandran plot						
Core (%)	191 (99.0%)	193 (99.5%)	189 (97.9%)	189 (97.4%)	195 (98.5%)	
Allowed (%)	2 (1.0%)	1 (0.5%)	4 (2.1%)	4 (2.1%)	3 (1.5%)	
Disallowed (%)	0 (0.0%)	0 (0%)	0 (0.0%)	1 (0.5%)	0 (0.0%)	
	4M9K	4M9M	4M9I	4M9F	4M9T	

[a]Data in parentheses are for the highest resolution bin.

Supplemental Figures

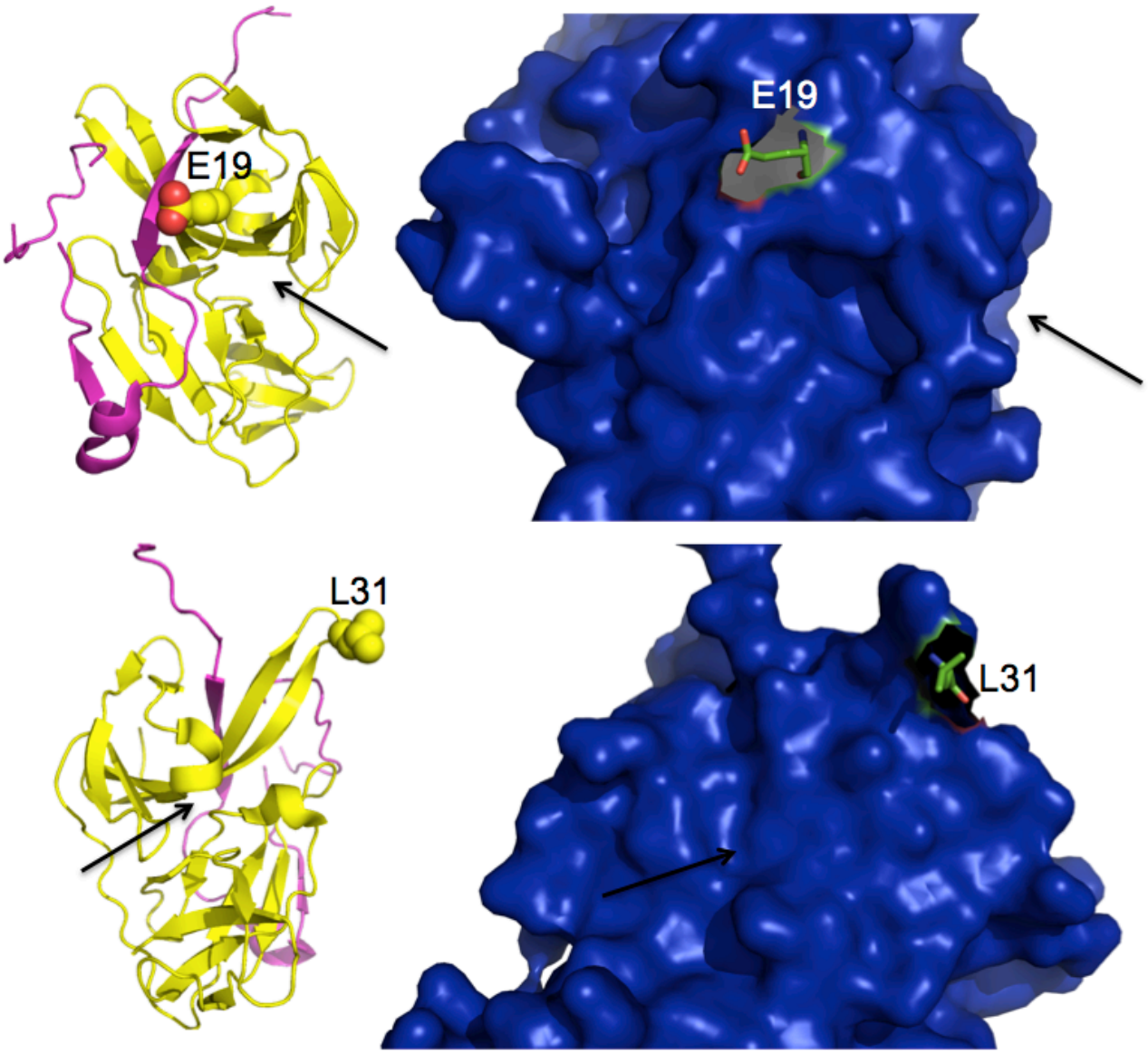


Figure S1A. Surface contours near cysteine substitutions. Putative allosteric sites were predicted by visual inspection of the published DENV2 NS2B-NS3pro crystal structure (pdb: 2FOM). Cysteines were introduced in proximity to cavities that were identified on the surface of NS3pro. Residues replaced by cysteine are shown as spheres in the ribbon diagram, which provides the orientation of the protein. In the surface respresentation (blue), the replaced residue is shown as green sticks. Among cysteine variants W83C, T111C and L115C did not show any catalytic activity while the others were as active as wild type NS2B-NS3pro. Arrows indicate the active site of NS2B-NS3pro.

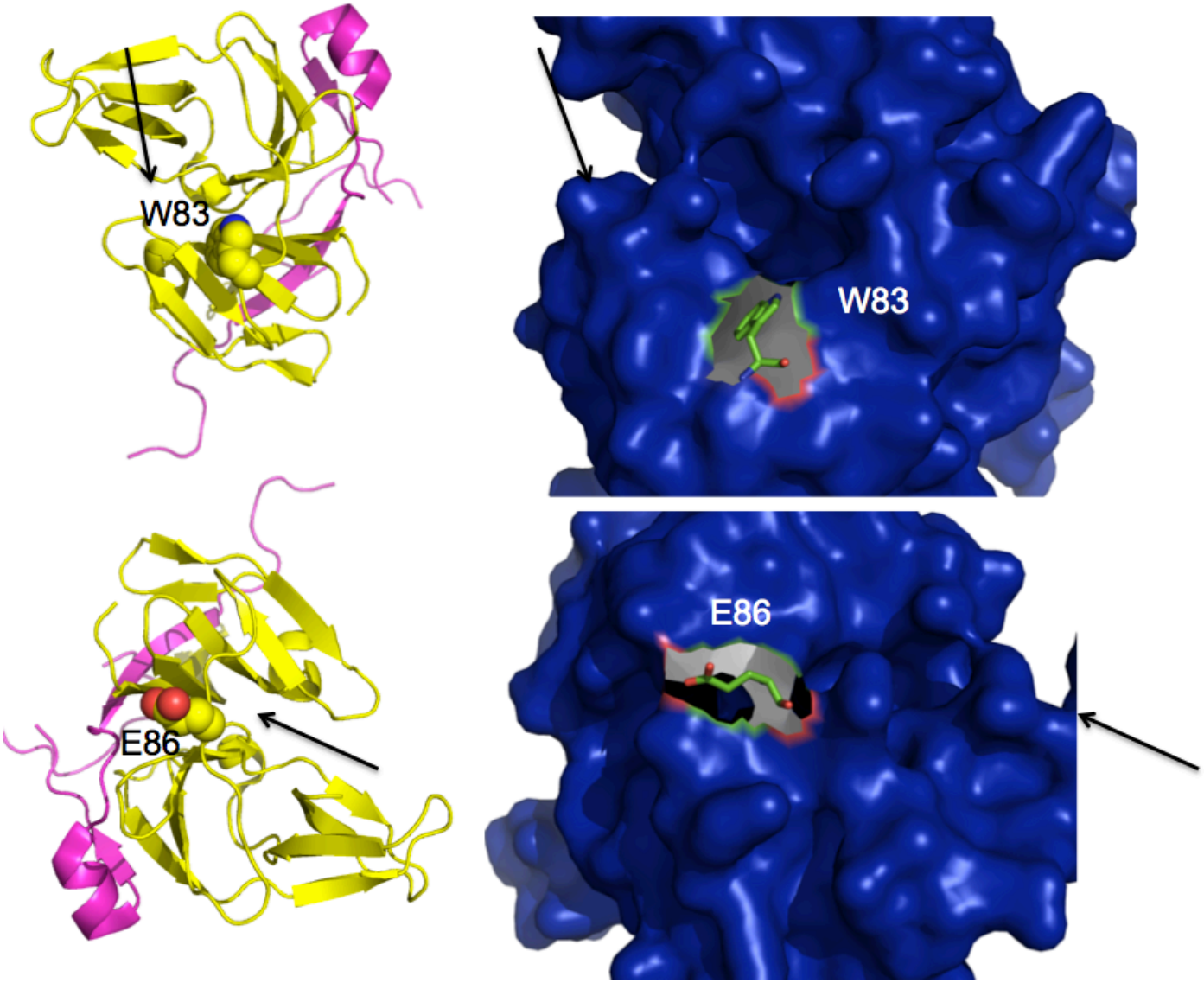


Figure S1B. Surface contours near cysteine substitutions. Putative allosteric sites were predicted by visual inspection of the published DENV2 NS2B-NS3pro crystal structure (pdb: 2FOM). Cysteines were introduced in proximity to cavities that were identified on the surface of NS3pro. Residues replaced by cysteine are shown as spheres in the ribbon diagram, which provides the orientation of the protein. In the surface respresentation (blue), the replaced residue is shown as green sticks. Among cysteine variants W83C, T111C and L115C did not show any catalytic activity while the others were as active as wild type NS2B-NS3pro. Arrows indicate the active site of NS2B-NS3pro.

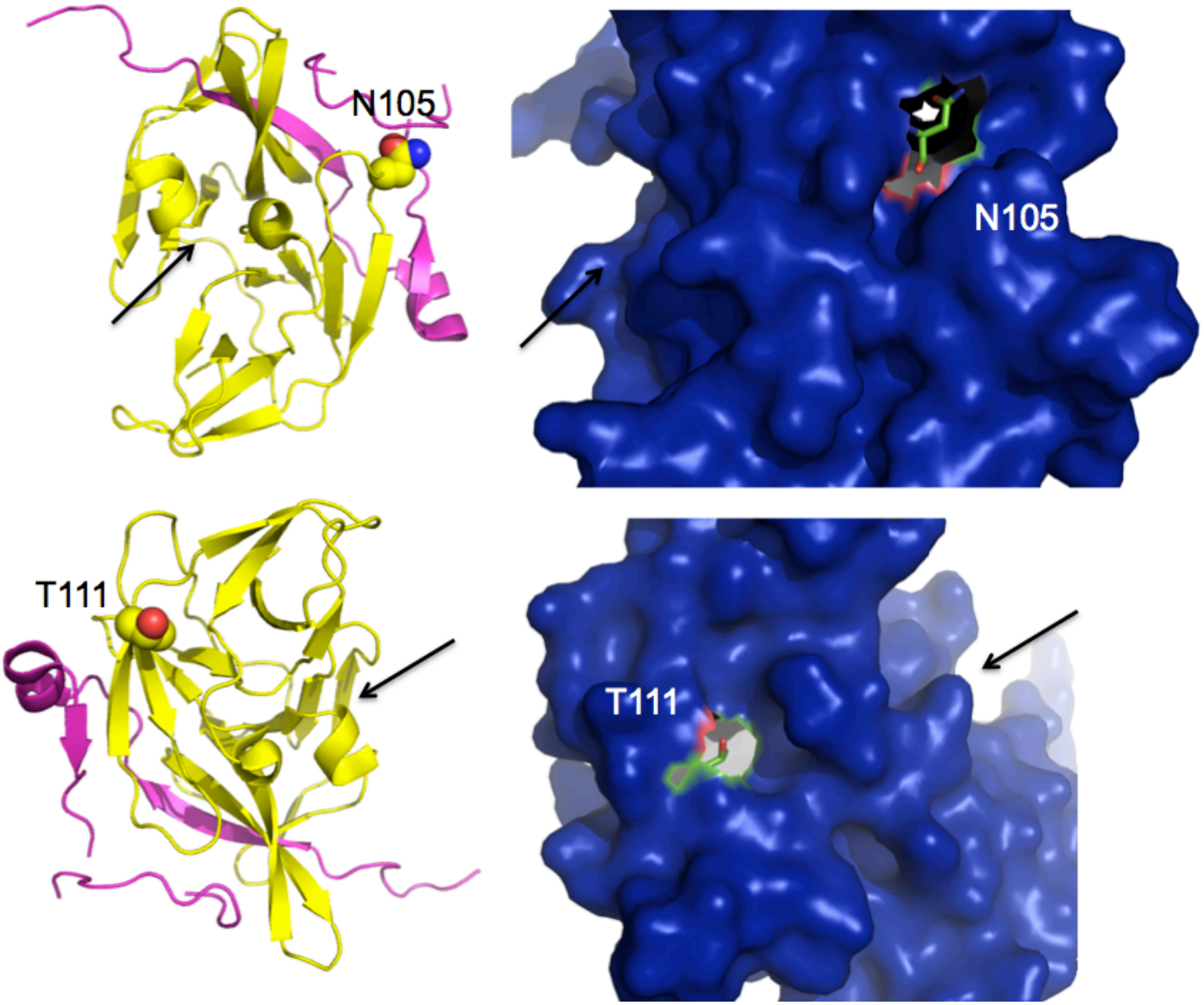


Figure S1C. Surface contours near cysteine substitutions. Putative allosteric sites were predicted by visual inspection of the published DENV2 NS2B-NS3pro crystal structure (pdb: 2FOM). Cysteines were introduced in proximity to cavities that were identified on the surface of NS3pro. Residues replaced by cysteine are shown as spheres in the ribbon diagram, which provides the orientation of the protein. In the surface respresentation (blue), the replaced residue is shown as green sticks. Among cysteine variants W83C, T111C and L115C did not show any catalytic activity while the others were as active as wild type NS2B-NS3pro. Arrows indicate the active site of NS2B-NS3pro.

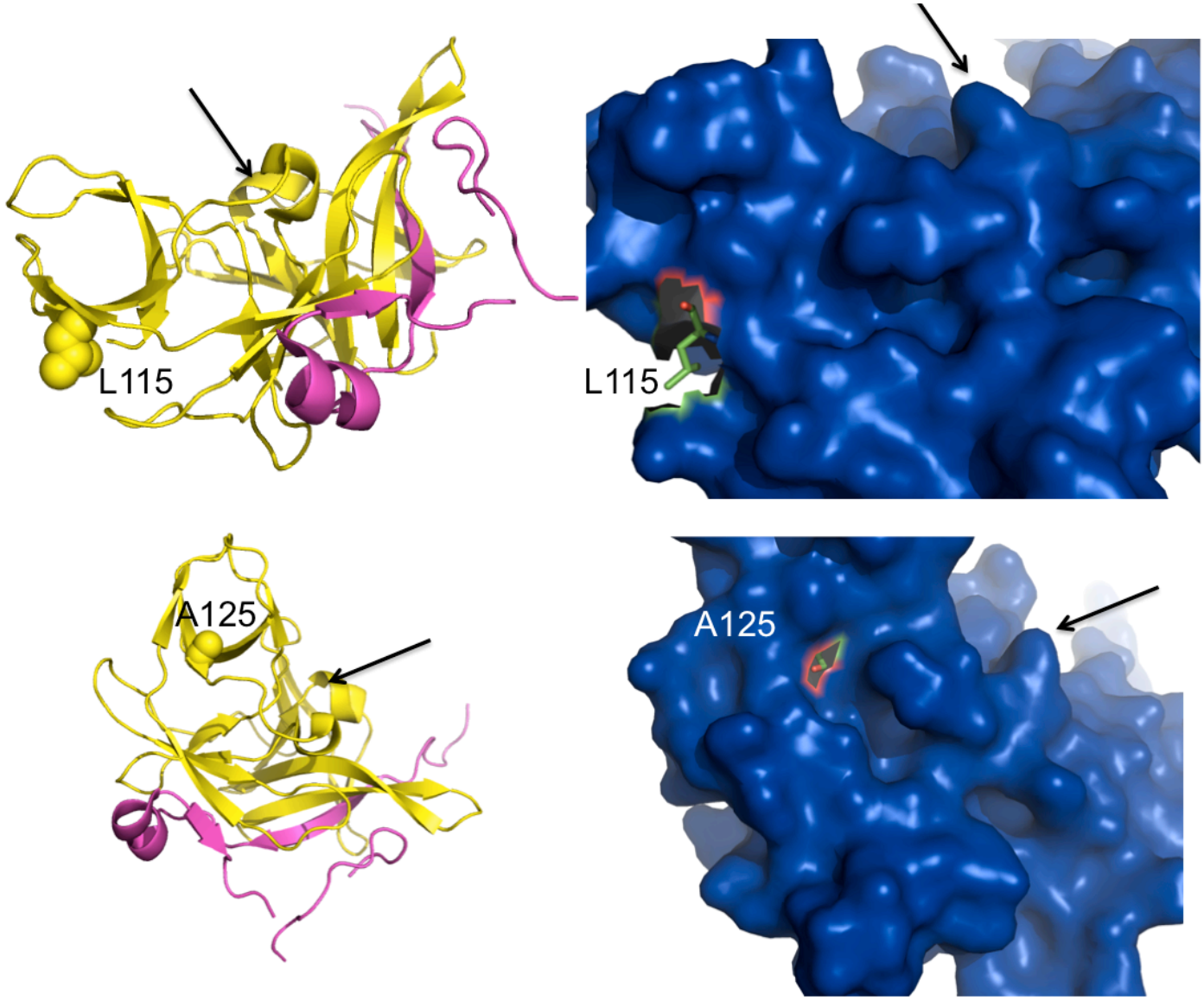
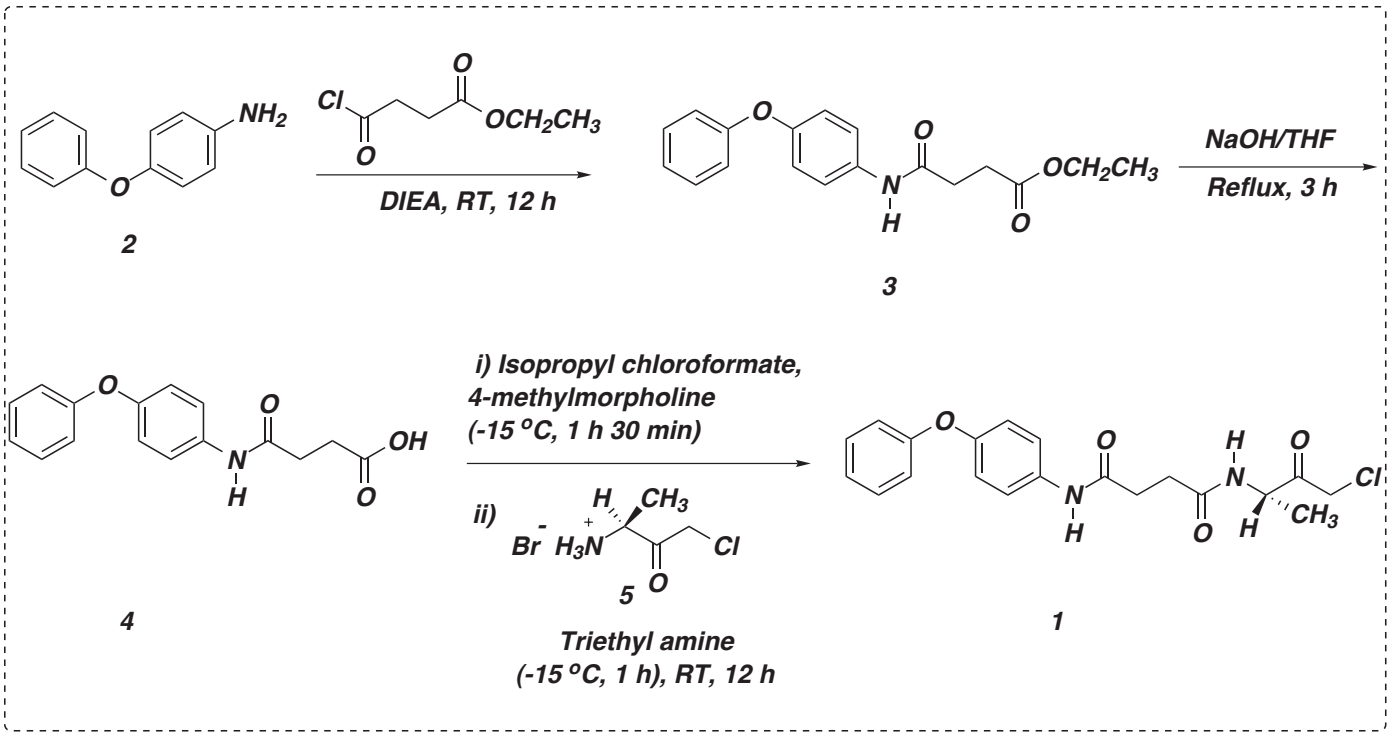


Figure S1D. Surface contours near cysteine substitutions. Putative allosteric sites were predicted by visual inspection of the published DENV2 NS2B-NS3pro crystal structure (pdb: 2FOM). Cysteines were introduced in proximity to cavities that were identified on the surface of NS3pro. Residues replaced by cysteine are shown as spheres in the ribbon diagram, which provides the orientation of the protein. In the surface respresentation (blue), the replaced residue is shown as green sticks. Among cysteine variants W83C, T111C and L115C did not show any catalytic activity while the others were as active as wild type NS2B-NS3pro. Arrows indicate the active site of NS2B-NS3pro.

A



B

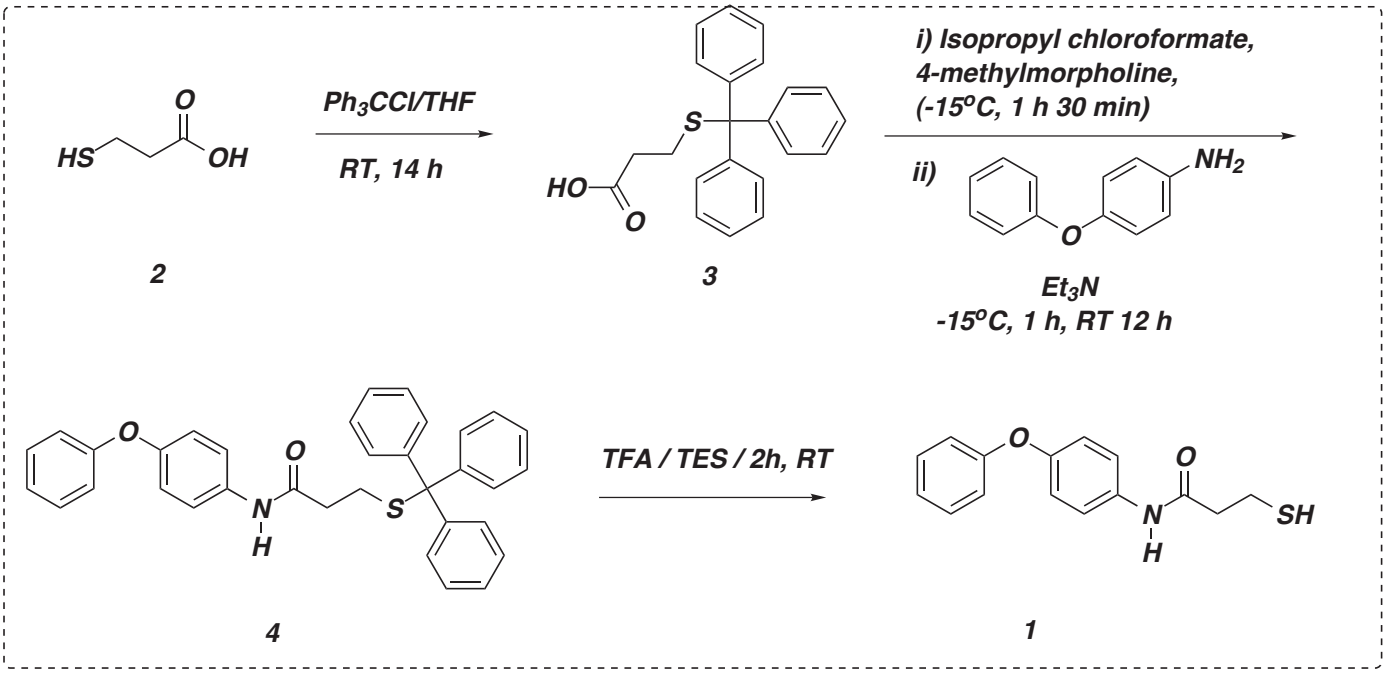


Figure S2. Synthetic scheme for production of biarylchloromethylketone and biarylthiol.

(A) The biarychlormethylketone (BAClMK) is a novel cysteine-reactive compound that is composed of a broadly-specific biarylether monophore attached to a highly reactive electrophilic chlormethylketone warhead via a linker. (B) A cysteine-reactive compound that contains the same broadly-specific hydrophobic monophore (biarylether), a linker and a thiol group for disulfide exchange.

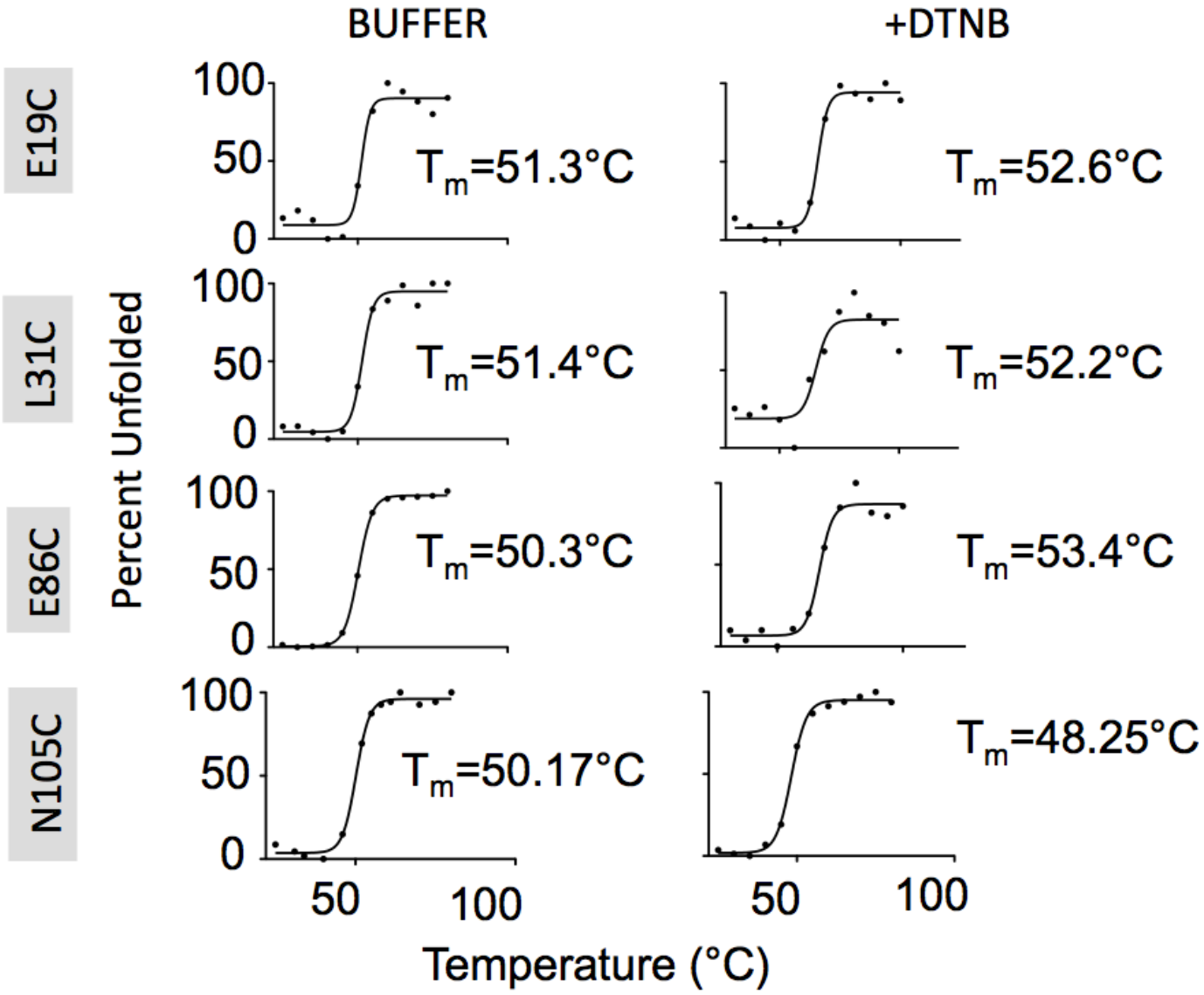


Figure S3. DTNB induces minimal changes in thermal stability.

The thermal stability as assessed by intrinsic tryptophan fluorescence of NS2B-NS3pro E19C, L31C, E86C and N105C in the presence or absence of the inhibitor DTNB are similar to eachother and to WT and A125C (Fig. 4), suggesting that neither cysteine substitution nor compound binding have any significant effect on stability or on the folded/unfolded equilibrium for the protein.

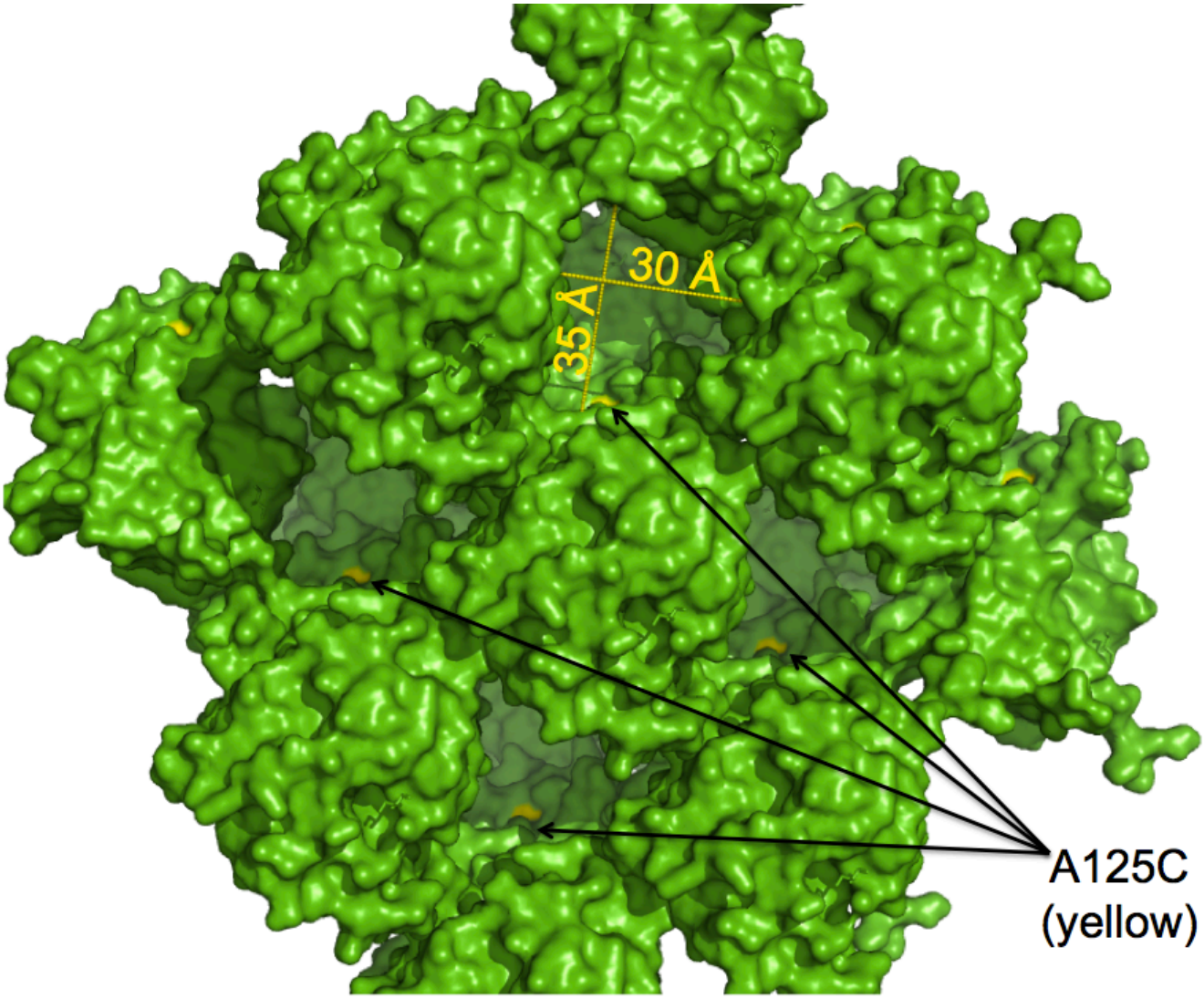


Figure S4. Solvent channels in the A125C crystals allow access of DTNB to C125.

The crystallographic packing of NS2B-NS3pro A125C features large solvent channels, approximately 30Å in diameter, observable as cavities between protein monomers. The introduced cysteine (yellow spots) are readily accessible inside these solvent channels.

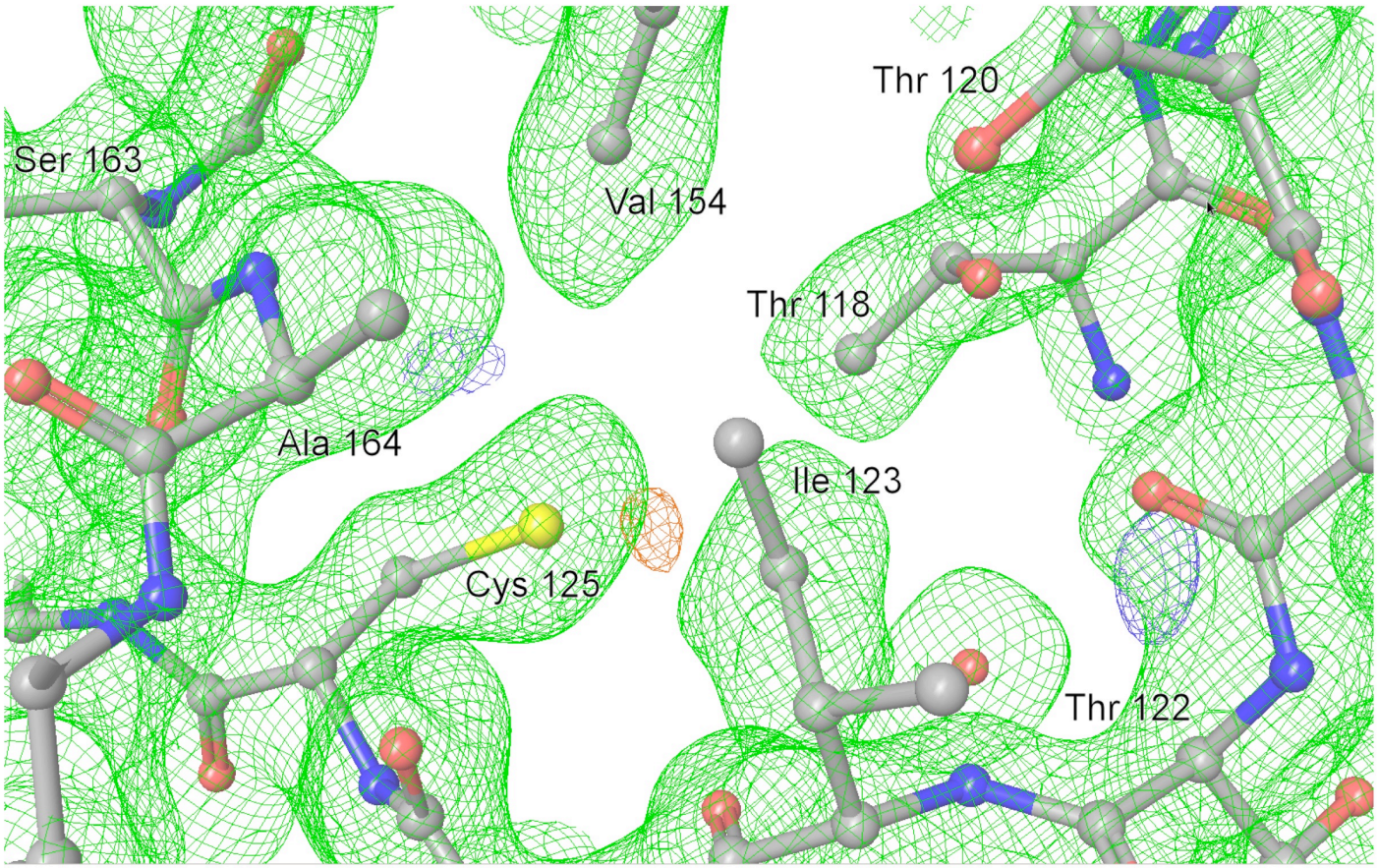


Figure S5. Region surrounding residue Cys 125 in the structure of A125C + DTNB, pH 8.5. All residues in this view are from the B chain. Hydrogen atoms have been omitted for clarity. Green contours are from a 2Fo-Fc map contoured at 1.0 σ, showing the electron density calculated from the observed diffraction data. The red and blue contours are from an Fo-Fc map contoured at 3.0 sigma, indicating a possible excess and deficiency of electrons in the atomic model, respectively. The Fo-Fc contours shown in this view are close to the noise level. The absence of any large Fo-Fc peaks indicates that the model fits the observed data well. No indication of DTNB binding can be seen in the crystallographic results, even though a chemical reaction with DTNB was clearly observed both in the crystal and in solution, and even though the addition of DTNB to the crystals changed the cell constants and the observed protein structure as described in the main text.

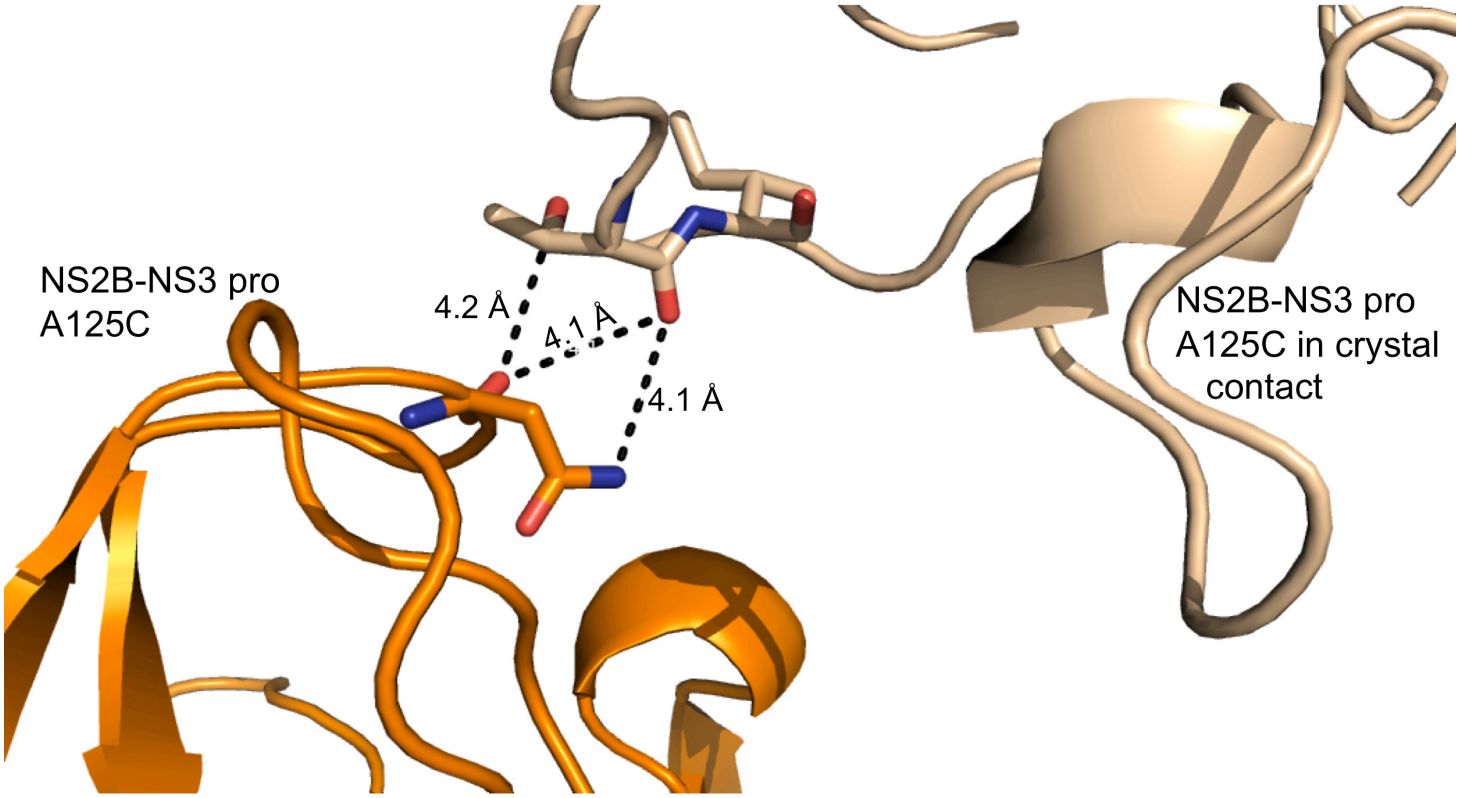


Figure S6. Crystal contacts between the 120s loop and NS2B in the presence of DTNB.

One crystal contact between N119 in the 120s loop and T77 in NS2B appear to be responsible for maintaining the observed conformation of the 120s loop.

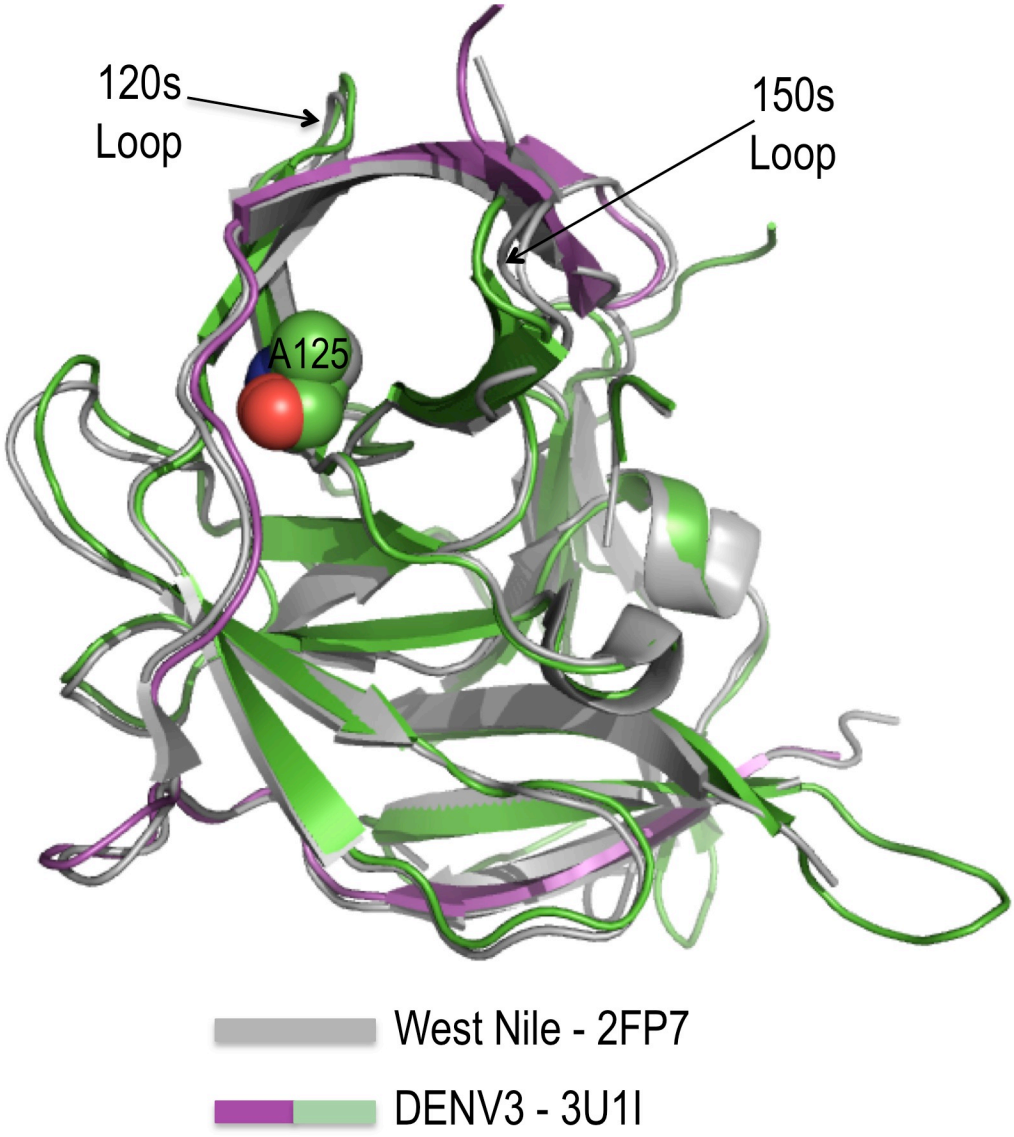


Figure S7. The A125 site is conserved in the NS2B-NS3 proteases from dengue virus and West Nile virus.

The A125-adjacent site that is susceptible to inhibition by DTNB, BAClMK and Aldrithiol lies between the 120s and 150s loops in DENV NS2B-NS3pro (purple and green strands) and in West Nile virus NS2B-NS3pro (gray).

*Thrombin cleavage site

Figure S8. Domain and construct architecture for dengue virus protease NS2B-NS3pro.

(A) The dengue virus protease construct used in these studies encodes the NS2B cofactor region (purple) covalently linked to NS3pro (yellow) via a Gly_4-Ser-Gly_4 linker. For purification purposes a His_6-tag (gray) was linked to NS2B via a thrombin cleavable site (TCS; black).