Genetic encoding of fluorinated analogues of phenylalanine for ^{19}F-NMR spectroscopy: detection of conformational heterogeneity in flaviviral NS2B-NS3 proteases

Haocheng Qianzhu,[1] Yi Jiun Tan,[2] Elwy H. Abdelkader,[1] Thomas Huber,[2,*] and Gottfried Otting[1,*]

[1] Research School of Chemistry, Australian National University, Canberra, ACT 2601, Australia

[2] ARC Centre of Excellence for Innovations in Peptide & Protein Science, Research School of Chemistry, Australian National University, Canberra, ACT 2601, Australia

Table of contents:

Figure S9. Intact protein mass spectra of eZika with 2F-Phe, 3F-Phe, 4F-Phe, 3,5diF-Phe, and 2,6diF-Phe in position 116

Supplementary tables

Table S1. Crystal structures of flaviviral NS2B-NS3 proteases in the protein data bank, where the constructs include the residues of NS2Bc

Table S2. Mutations found in identified G1PylRS variants that recognize 2F-Phe, 4F-Phe, and F_5-Phe

Table S3. Protein yields of NS2B-NS3 proteases with fluorinated phenylalanines in position 116

Table S4. DNA and corresponding amino acid sequences of the proteins used

References

a) Materials

The non-canonical amino acids (ncAAs) 2-fluoro-L-phenylalanine (2F-Phe, Cat. No.: A191400), 3-fluoro-L-phenylalanine (3F-Phe, Cat. No.: A327766), 2,6-difluoro-L-phenylalanine (2,6diF-Phe, Cat. No.: A395183), 3,5-difluoro-L-phenylalanine (3,5diF-Phe, Cat. No.: A131886), and 2,3,4,5,6-pentafluoro-L-phenylalanine (F_5-Phe, Cat. No.: A142712) were purchased from Ambeed (USA) and used without further purification. 4-Fluoro-L-phenylalanine (4F-Phe, Cat. No.: J53633) was purchased from AK Scientific Inc. (USA) and used without further purification.

b) Selection of functional G1PylRS enzymes recognizing fluorinated phenylalanines

The selection of functional G1PylRS enzymes followed a previously established protocol, using the previously established library of G1PylRS mutants encoded on the pBK-G1RS plasmid.[1-2] The plasmid library was transformed into *E. coli* DH10B cells harboring the pBAD-H6RFP reporter plasmid encoding mCherry red fluorescent protein (RFP) with an amber stop codon following an N-terminal His_6 tag. After transformation, the culture was directly inoculated into 25 mL LB medium supplemented with 100 mg/L carbenicillin, 50 mg/L kanamycin, 0.4% L-arabinose, and 1 mM target ncAA. This culture served as the sample for the first round of positive selection (**1P+**). The same cells grown without the supplement of ncAA were used as a control sample (**1P–**). Overnight expression at 37 °C led to readily detectable level of RFP expression, and then, the cells were harvested, resuspended in 5 mL of PBS buffer (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, pH 7.4) and diluted 100-fold to a concentration suitable for fluorescence-activated cell sorting (FACS). FACS was performed using a FACSAria Fusion cell sorter (BD Biosciences, USA; Figures S2, S4, and S5). Cells with high red fluorescence levels were collected from the **1P+** sample (indicated by violet shades in Figures S2b, S4a, and S5a) and subjected to a subsequent round of negative selection (**2N–**) in the absence of ncAA. Cells exhibiting low RFP expression were collected from the **2N–** sample and aliquoted to inoculate media under positive (**3P+**, with ncAA) and negative (**3P–**, without ncAA) conditions. The **3P+** sample for F_5Phe demonstrated a clear response to the presence of F_5Phe, and the top 5.1% of RFP-fluorescent cells from the **3P+** sample were collected. The selection experiments for 2F-Phe and 4F-Phe continued to the fourth and fifth round following the same strategy. Both reached enrichment on the target population. Approximately the top 3% of RFP-fluorescent cells from the **5P+** samples of each experiment were collected.

About 2,000 cells collected from each final round were recovered by plating on LB agar plates containing 100 mg/L carbenicillin and 50 mg/L kanamycin. Isolated colonies were analyzed in 96-well plates. Sixty enzyme candidates were inoculated into media under both positive (with 1 mM ncAA) and negative (without ncAA) conditions. The red fluorescence intensity was measured as an indicator of RFP expression and normalized to the OD_{600} of the cell culture using a TECAN Infinite 200 Pro M Plex plate reader (Tecan, Switzerland; Figures S2c, S4b, and S5b). The DNA sequence analysis of the pBK-G1RS plasmids identified six, three, and four individually different candidates to incorporate 2F-Phe, 4F-Phe, and F_5-Phe, respectively. The amino acid mutations of these candidates are listed in Table S1. Mutant G1-F5F29 was found to be polyspecific towards 3F-Phe, 2,6diF-Phe, and 3,5diF-Phe and used for all applications installing these ncAAs.

c) Protein expression and purification

The genes of G1-2FF44, G1-4FF02, and G1-F5F29 were cloned into a high-copy number pRSF plasmid to yield the new tRNA synthetase plasmids pRSF-G1-2FF44/4FF02/F5F29. The pCDF plasmids containing the gene of the protein of interest interrupted by an amber codon were transformed into *E. coli* B-95.ΔAΔ*fabR* cells[3] together with the appropriate pRSF plasmid. The transformed cells were grown at 37 °C in LB medium containing 25 mg/L kanamycin and 25 mg/L spectinomycin. A 10 mL overnight culture was used to inoculate 1 L LB medium supplemented with 25 mg/L kanamycin and 25 mg/L spectinomycin. The cells were grown at 37 °C to an OD_{600} of 0.5, at which point the requisite fluorinated phenylalanine analogue was added in a final concentration of 1 mM. Subsequently, the temperature was reduced to 25 °C and protein expression was induced by the addition of 1 mM IPTG.

After expression for 16 h, the cells were harvested by centrifugation. Following resuspension in buffer A (50 mM Tris-HCl pH 7.5, 300 mM NaCl), the cells were lysed using an Avestin Emulsiflex C5 system (Avestin, Canada) using two passes with a pressure of 10,000–15,000 psi. The cell lysates were centrifuged for 1 h at 30,000 *g*. The supernatant was loaded onto a 1 mL His GraviTrap column (Cytiva, USA). The column was washed with 20 column volumes buffer A and the protein was eluted with 5 column volumes buffer B (same as buffer A but with 300 mM imidazole). Analysis of the eDENV and eZika constructs by SDS-PAGE and mass spectrometry showed that self-cleavage was complete after purification. For NMR measurements, the buffer was exchanged to 50 mM HEPES pH 7.5 and 100 mM NaCl, and 10% D_2O and 0.1 mM trifluoroacetic acid (TFA) were added. To produce the complexes with BPTI, BPTI was added from a 7.7 mM stock solution in water. For the complexes with inhibitor **1**, the inhibitor was added using a 100 mM stock solution in DMSO.

d) Protein mass spectrometry

Intact protein analysis was performed on an Orbitrap FusionTM TribridTM mass spectrometer (Thermo Fisher Scientific, USA) connected to a Thermo Fisher Scientific UltiMate 3000 HPLC system equipped with ZORBAX 300SB-C3, 3.5 μm, 4.6 x 50 mm HPLC column (Agilent Technologies, USA). Approximately 50 pmol of sample was injected using a 500 μL/min linear gradient of solvent A (0.1% (v/v) formic acid in water) and solvent B (0.1% (v/v) formic acid in acetonitrile), ramping solvent B from 5% solvent B at the start to 80% after 12 min. Data were collected using an electrospray ionization (ESI) source in positive ion mode. Protein intact mass was determined by deconvolution using the program Xcalibur 3.0.63 (Thermo Fisher Scientific, USA).

e) NMR spectroscopy

^{19}F NMR spectra were recorded at 25 ^{o}C without ^{1}H decoupling on either a Bruker 600 MHz NMR spectrometer equipped with a 5 mm TCI cryoprobe with the inner coil tuned to ^{19}F or a Bruker 400 MHz NMR spectrometer equipped with a broadband probe. Samples were in 5 mm NMR tubes at 0.2 – 1.2 mM concentrations (DENV) or 0.1 – 0.7 mM (Zika). Samples with concentrations above 0.4 mM were diluted to 0.2 mM prior the addition of BPTI or inhibitor **1**. Parameters used: acquisition time 120 ms (600 MHz NMR spectrometer) or 218 ms (400 MHz NMR spectrometer), exponential window multiplication with line broadening by 30 – 100 Hz prior to Fourier transformation. Typical recording times of the NMR spectra were about 1 to 3 hours. The only exception was a second preparation of gDENV with 3F-Phe to probe the complex with inhibitor **1**, for which the spectra were recorded overnight. On the 600 MHz NMR spectrometer, measurement times were as short as 15 minutes. The ^{19}F chemical shifts were calibrated relative to internal TFA (-75.25 ppm).

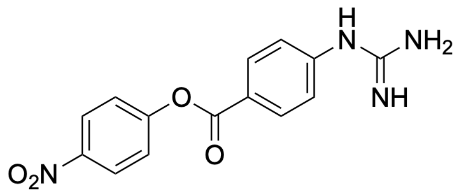

Figure S1. Chemical structure of inhibitor **1**. The compound forms an ester with the active site serine residue of serine proteases with the release of yellow *p*-nitrobenzoate.

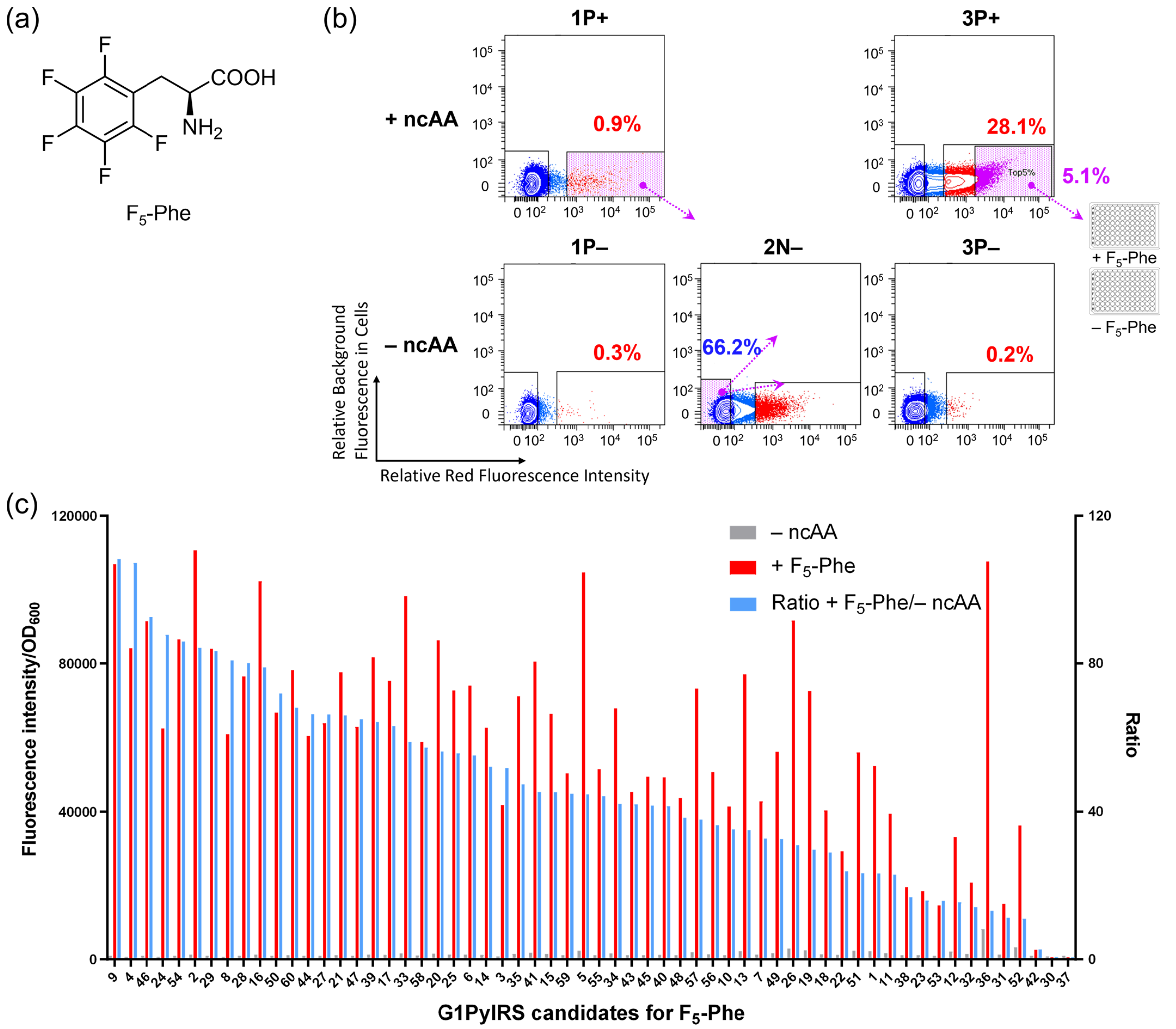

Figure S2. FACS experiments for selecting functional G1PylRS enzymes for F$_5$-Phe. (a) Chemical structure of F$_5$-Phe. (b) FACS screening of G1PylRS variants for activity and specificity in recognizing F$_5$-Phe. The horizontal axis of the scatter plots represents red fluorescence intensity (excitation at 560 nm), while the vertical axis indicates background fluorescence in cells excited at 488 nm. Positive (**P**) and negative (**N**) selection rounds are indicated with "+", denoting growth conditions with 1 mM ncAA, and "–" for conditions without ncAA. Violet-shaded regions identify the collected cell populations. Arrows illustrate the subsequent selection strategy applied after amplification by culturing. (c) Activity and specificity screen of G1PylRS variants for F$_5$-Phe incorporation. Cells from the 3.1% fraction with the highest red fluorescence in the final selection round were cultured in 96-well plates with and without 1 mM F$_5$-Phe. Red fluorescence intensity indicative of the readthrough efficiency of the amber-interrupted reporter gene was then measured. The plot presents the colonies ranked in a descending order based on the ratio of red fluorescence in the + ncAA wells compared to the – ncAA wells. This ranking highlights the candidates with the highest activity and specificity for F$_5$-Phe incorporation.

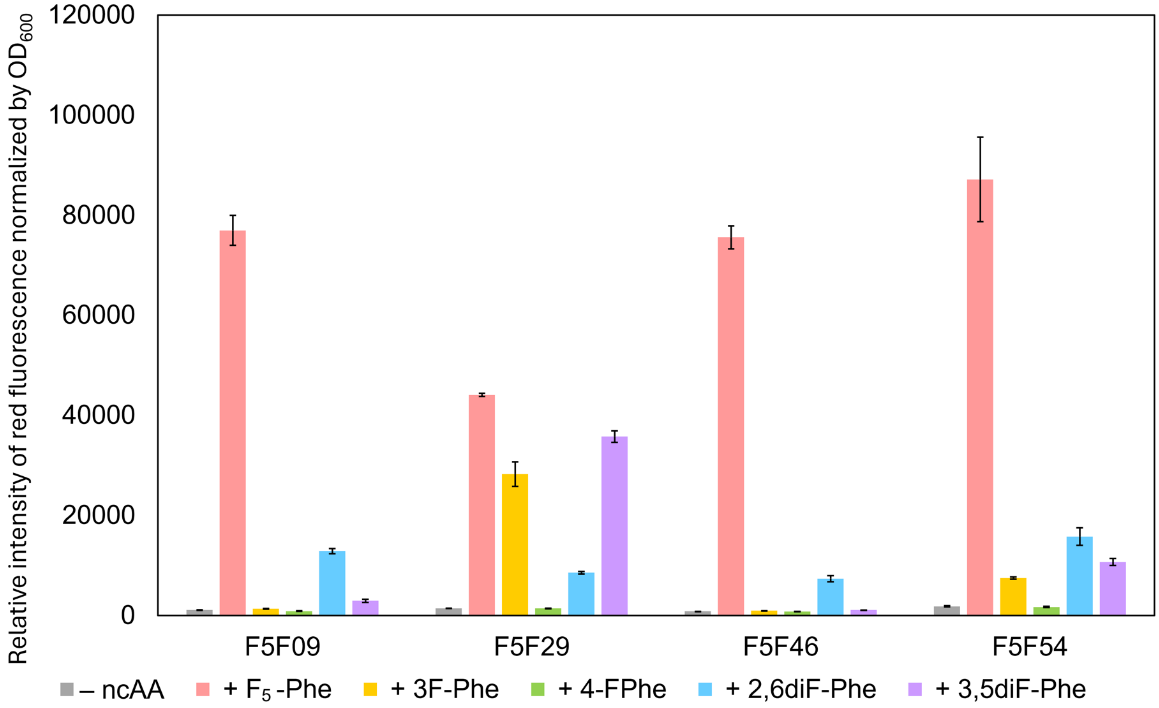

Figure S3. Substrate polyspecificity of the enzymes F5F09, F5F29, F5F46, and F5F54 (Table S2) selected for F_5-Phe towards different fluorinated phenylalanines. To indicate the enzyme activity, levels of fluorescence of amber-interrupted RFP were measured in a plate reader, normalized by OD, and averaged across biological triplicates, following bacterial growth in the presence of 1 mM of different fluorinated phenylalanines or without addition of any ncAA. Standard deviation is indicated by the error bars.

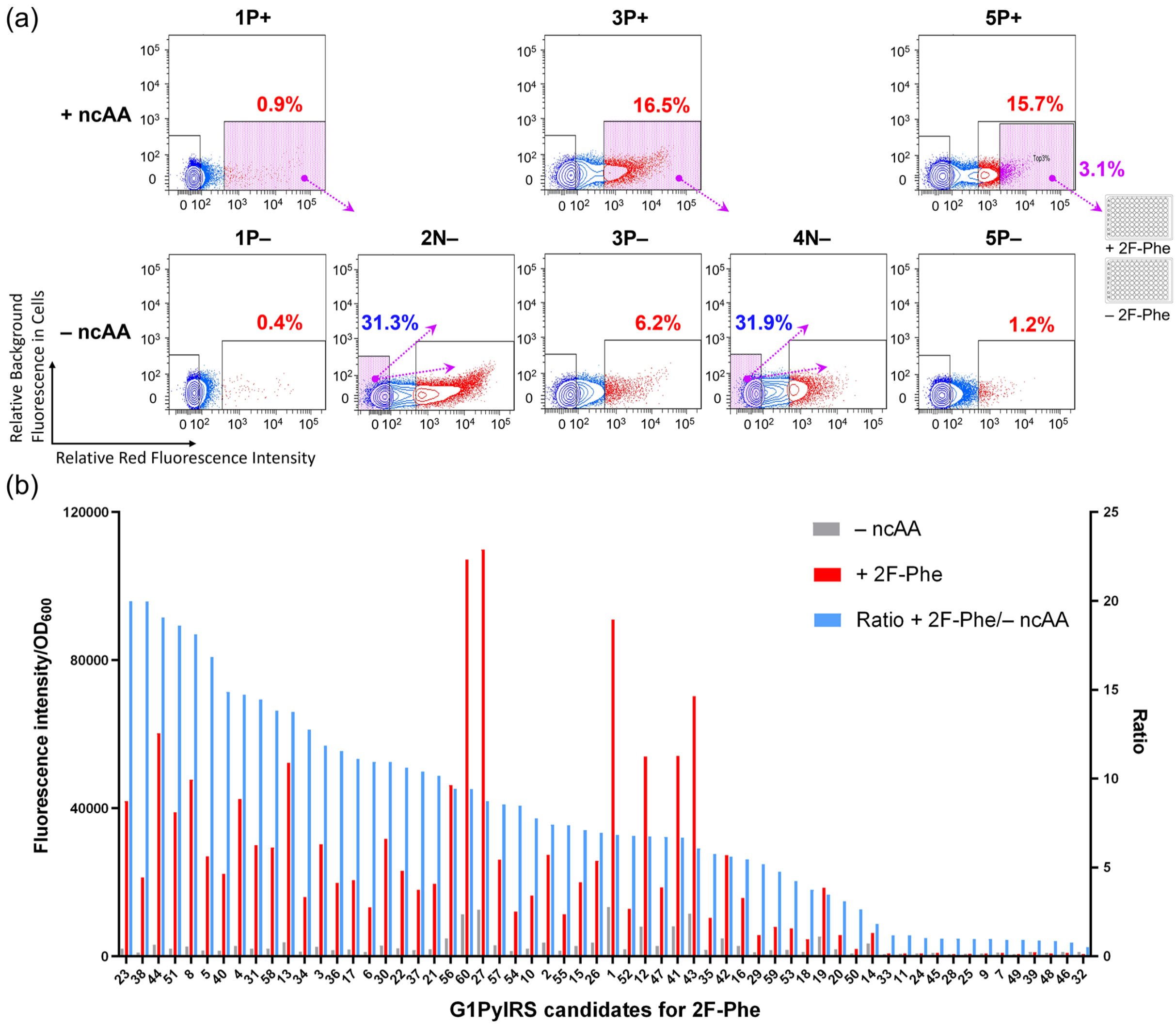

Figure S4. FACS experiments for selecting functional G1PylRS enzymes for 2F-Phe. (a) FACS screening of G1PylRS variants for activity and specificity in recognizing 2F-Phe. Annotations are the same as in Figure S1. (b) Activity and specificity screen of G1PylRS variants for 2F-Phe incorporation. Cells from the 5.1% fraction with the highest red fluorescence in the final selection round were cultured in 96-well plates with and without 1 mM 2F-Phe. Same as in Figure S1, the plot presents the colonies ranked in a descending order based on the ratio of red fluorescence in the + ncAA wells compared to the – ncAA wells, highlighting the candidates with the highest activity and specificity for 2F-Phe incorporation.

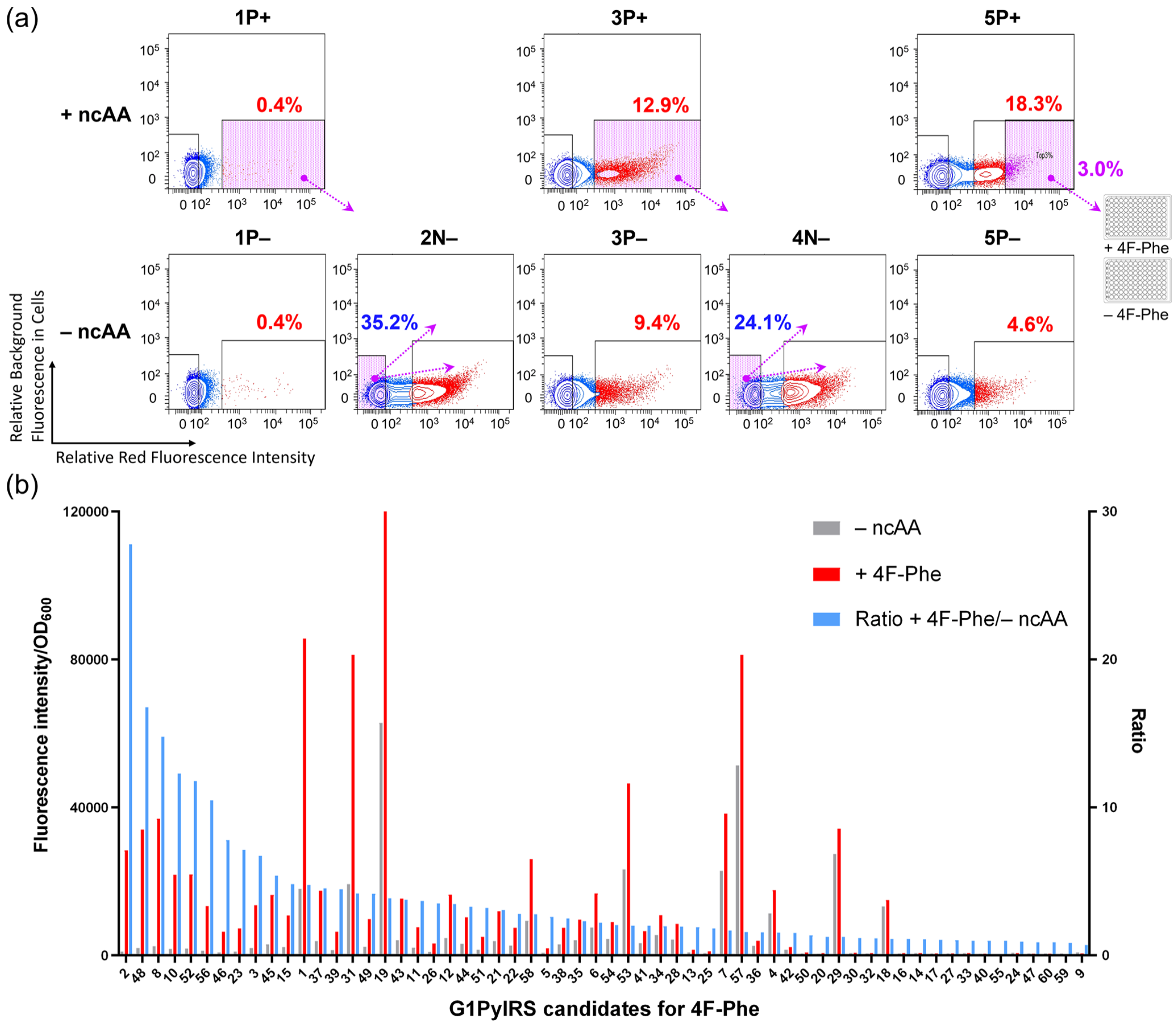

Figure S5. FACS experiments for selecting functional G1PylRS enzymes for 4F-Phe. (a) FACS screening of G1PylRS variants for activity and specificity in recognizing 4F-Phe. Annotations are the same as in Figure S1. The **1P–** control is the same sample as in the experiments for 2F-Phe. (b) Activity and specificity screen of G1PylRS variants for 4F-Phe incorporation. Cells from the 5.1% fraction with the highest red fluorescence in the final selection round were cultured in 96-well plates with and without 1 mM 4F-Phe. Same as in Figure S1, the plot presents the colonies ranked in a descending order based on the ratio of red fluorescence in the + ncAA wells compared to the – ncAA wells, highlighting the candidates with the highest activity and specificity for 4F-Phe incorporation.

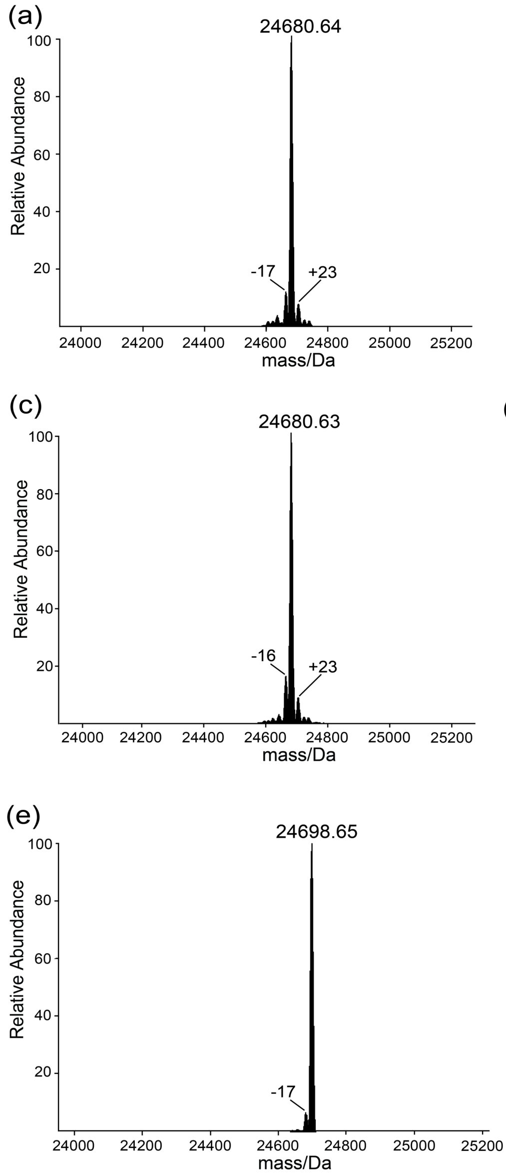

Figure S6. Intact protein mass spectra of gDENV with (a) 2F-Phe, (b) 3F-Phe, (c) 4F-Phe, (d) 3,5diF-Phe, and (e) 2,6diF-Phe in position 116. The calculated mass is 24,681.68 Da for the constructs containing singly fluorinated phenylalanine and 24,699.68 Da for the construct containing two fluorine atoms. Minor peaks are labeled with their mass relative to the mass of the main peak.

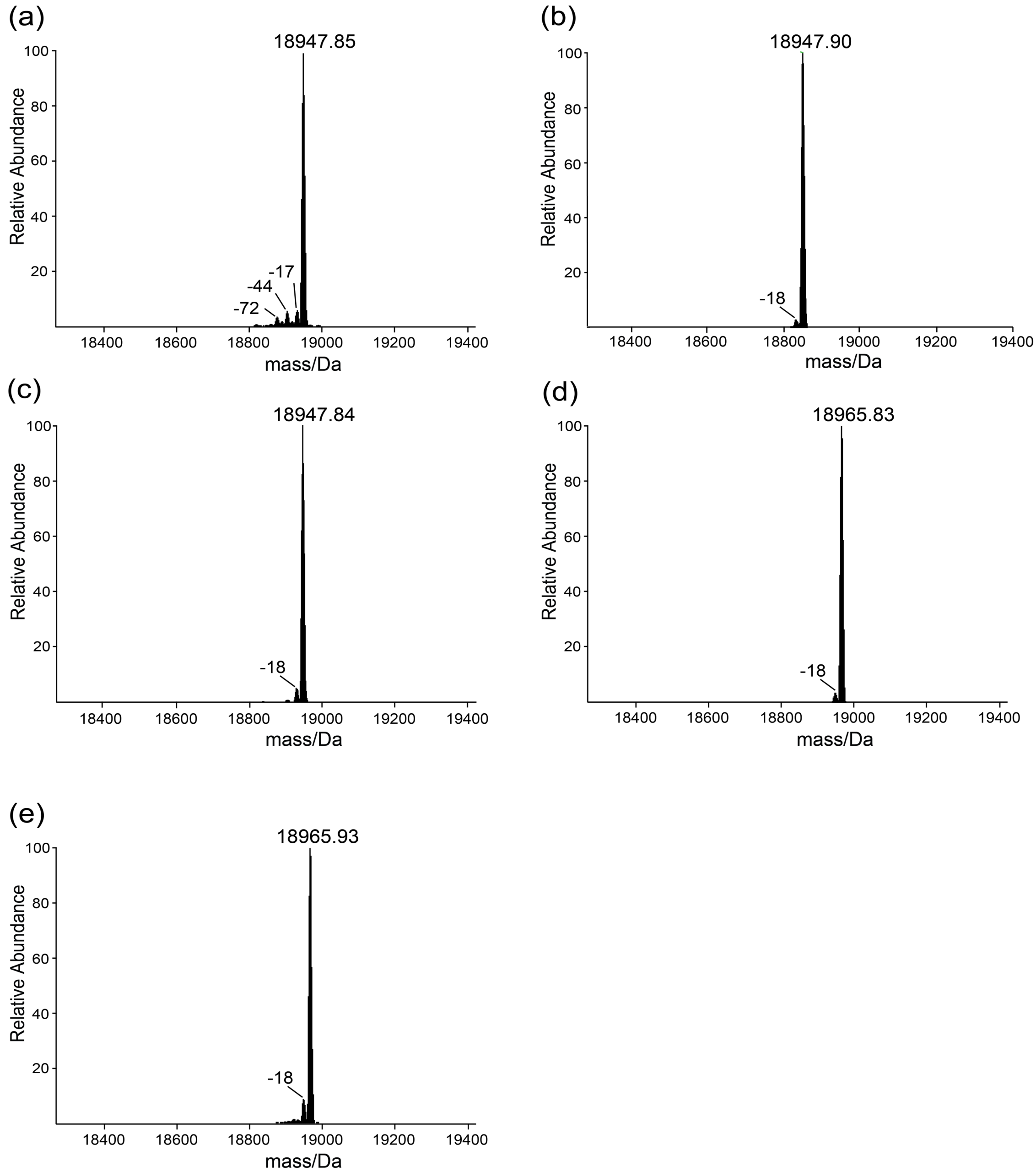

Figure S7. Intact protein mass spectra of the NS3 part of eDENV with (a) 2F-Phe, (b) 3F-Phe, (c) 4F-Phe, (d) 3,5diF-Phe, and (e) 2,6diF-Phe in position 116. The calculated mass is 18,948.58 Da for the constructs containing singly fluorinated phenylalanine and 18,966.58 Da for the construct containing two fluorine atoms. Minor peaks are labeled with their mass relative to the mass of the main peak.

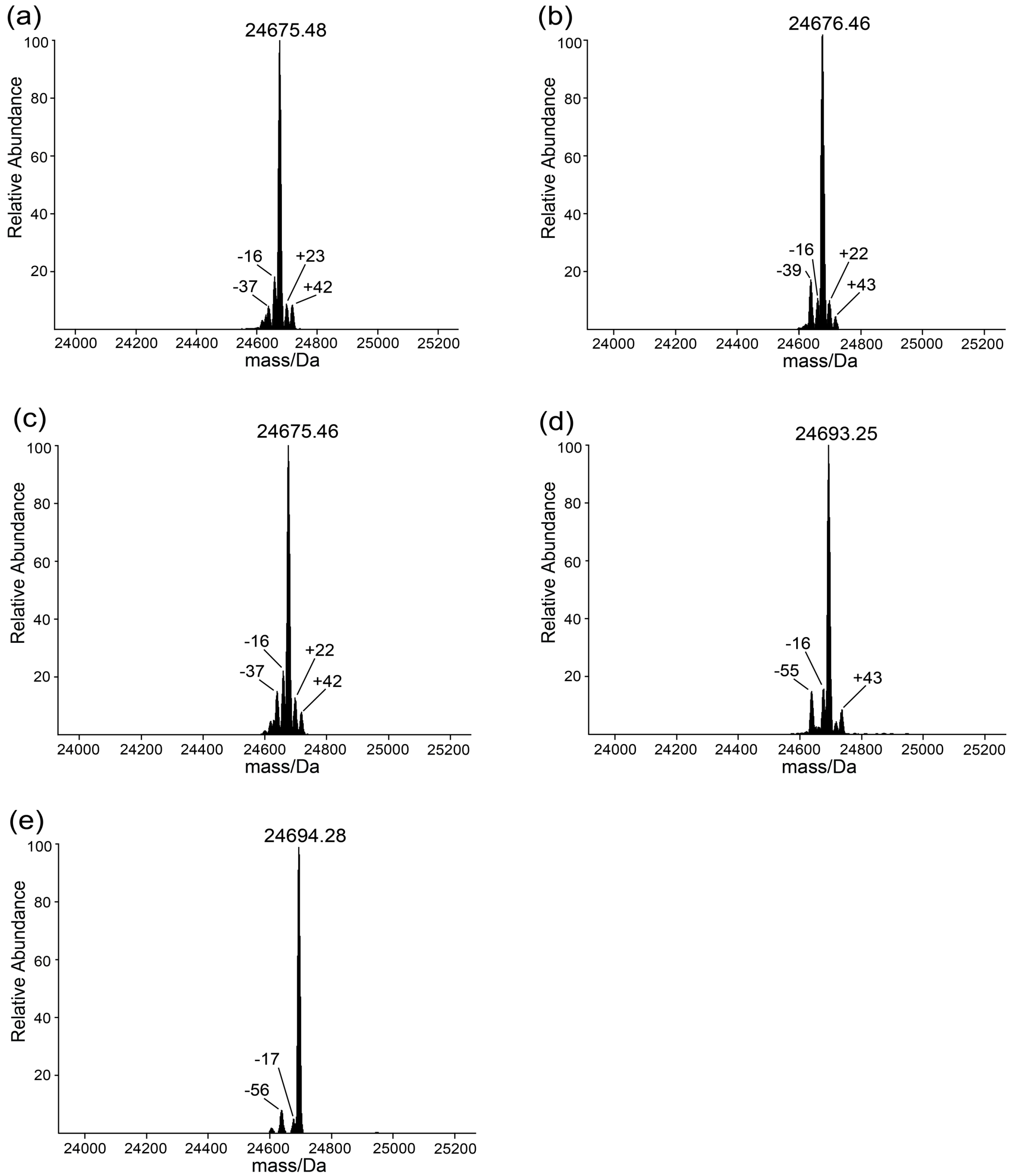

Figure S8. Intact protein mass spectra of gZika with (a) 2F-Phe, (b) 3F-Phe, (c) 4F-Phe, (d) 3,5diF-Phe, and (e) 2,6diF-Phe in position 116. The calculated mass is 24,676.44 Da for the constructs containing singly fluorinated phenylalanine and 24,694.44 Da for the constructs containing two fluorine atoms. Minor peaks are labeled with their mass relative to the mass of the main peak.

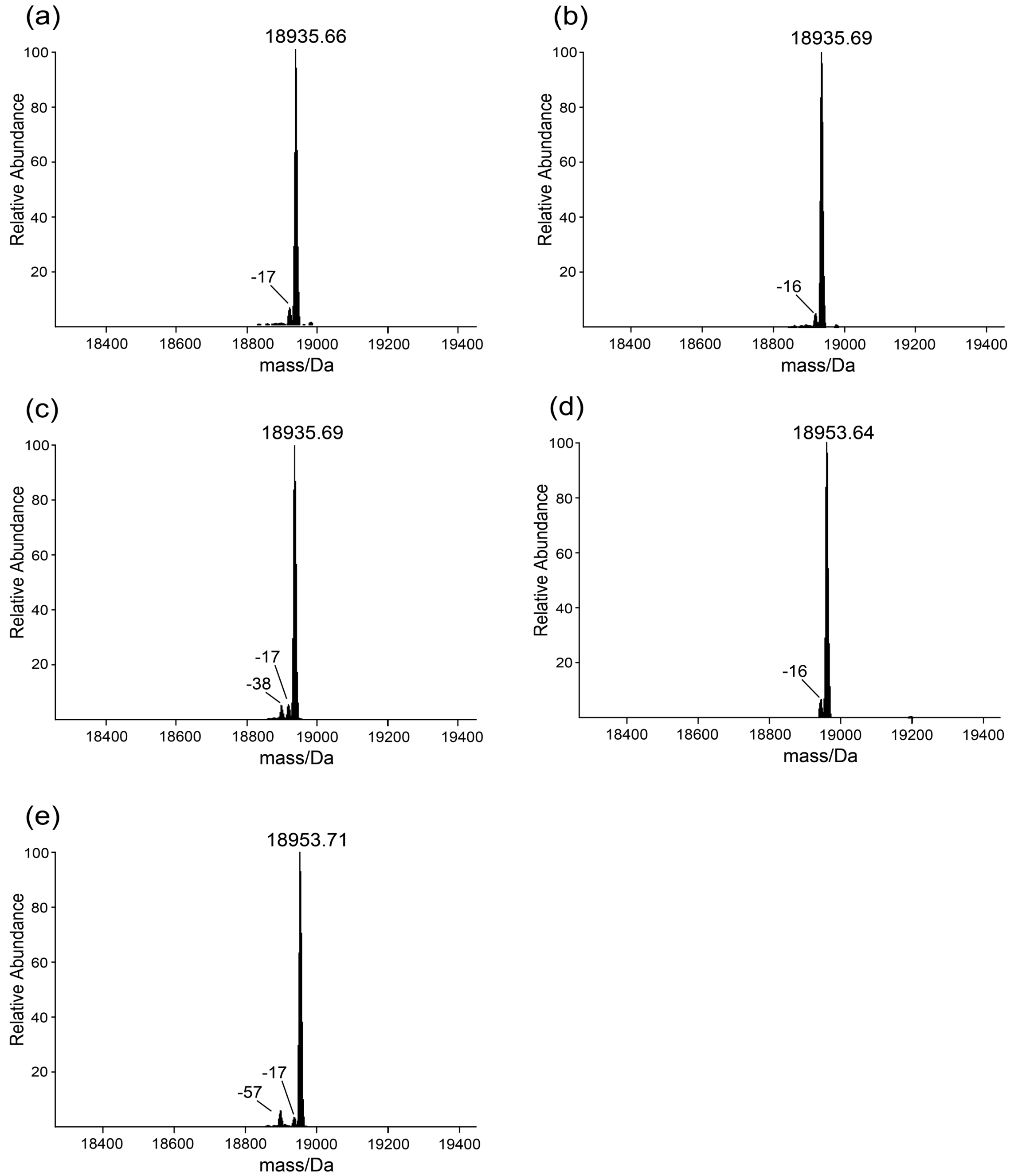

Figure S9. Intact protein mass spectra of the NS3 part of eZika with (a) 2F-Phe, (b) 3F-Phe, (c) 4F-Phe, (d) 3,5diF-Phe, and (e) 2,6diF-Phe in position 116. The calculated mass is 18,936.33 Da for the constructs containing singly fluorinated phenylalanine and 18,954.33 Da for the construct containing two fluorine atoms. Minor peaks are labeled with their mass relative to the mass of the main peak.

Table S1. Crystal structures of flaviviral NS2B-NS3 proteases in the protein data bank, where the constructs include the residues of NS2Bc.

Protein	NS2Bc conformation[a]	Phe116 side chain	Tryptophan side chain conformations[b]	PDB code	Resolution (Å)	Reference
DENV4 with full-length NS3 and G_4SG_4 linker	unobservable after E63*	Solvent-exposed, non-canonical orientation	canonical	5YVJ	2.5	[4]
DENV4 with full-length NS3, self-cleaved, 24 more residues in linker	closed	Not defined beyond C^{β}	canonical	5YVY	3.2	[4]
DENV4 with full-length NS3 and G_4SG_4 linker	closed	Not defined beyond C^{β}	W50 and W89 flipped, W83 undefined	5YVV	3.1	[4]
DENV4 with full-length NS3, G_4SG_4 linker	closed	Not defined beyond C^{β}	W69 and W89 flipped, W83 undefined	5YVW	3.1	[4]
DENV4 with full-length NS3 + BPTI, G_4SG_4 linker	closed	Close to F83*	W83 flipped	5YVU	2.5	[4]
DENV4 with full-length NS3 + BPTI, enzymatically cleaved	closed	Close to F83*, which has a χ_2 angle different from 5YVU	W69 and W83 flipped	5YW1	2.6	[4]
DENV4, self-cleaved	half-open	Buried but far from F83*	canonical	7VMV	3.35	[5]
DENV3 + inhibitor, G_4SG_4 linker	closed	M83* replaces F83* in this serotype. F116 has non-canonical ψ angle	canonical	3U1I	2.3	[6]
DENV3 + BPTI, G_4SG_4 linker	unobservable after G69*	Solvent-exposed	W50 flipped	3U1J	1.8	[6]
DENV2, G_4SG_4 linker	open	Solvent-exposed. In this serotype, F83* is replaced by M84*.	canonical	2FOM	1.5	[7]

DENV2, G_4SG_4 linker	open	Solvent-exposed. M84* in place of F83*	canonical	4M9F 4M9I 4M9K 4M9M 4M9T	1.5 2.4 1.5 1.5 1.7	[8]
DENV2 + allosteric inhibitor, G_4SG_4 linker	unobservable after S71*	Solvent-exposed	canonical	6MO0 6MO1 6MO2	2.7 3.0 2.8	[9]
DENV1, G_4SG_4 linker	open	Solvent-exposed. No aromatic residue equivalent to F83* in this serotype	canonical	3LKW 3L6P	2.0 2.2	[10]
WNV + inhibitor, self-cleaved	closed	Close to F85*	canonical	2FP7	1.7	[7]
WNV + BPTI, G_4SG_4 linker	closed	Close to F85*	canonical	2IJO	2.3	[11]
WNV, G_4SG_4 linker	open	Solvent-exposed	canonical	2GGV	1.8	[11]
WNV, G_4SG_4 linker	closed	Close to F85*	canonical	3E90	2.45	[12]
WNV + inhibitor, G_4SG_4 linker	closed	Close to F85*	canonical	2YOL	3.2	[13]
WNV + inhibitor, G_4SG_4 linker	closed	Close to F85*	W50 in two conformations	5IDK	1.5	[14]
WNV, G_4SG_4 linker	open	Solvent-exposed	canonical	8CO8	1.9	Fairhead et al., to be published
Zika + inhibitor, G_4SG_4 linker	closed	Close to F84*	W89 flipped in one of two conformations	5LC0	2.7	[15]
Zika bound to TGKR peptide, self-cleaved	closed	Close to F84*	canonical	5GJ4	1.8	[16]
Zika + inhibitor/ fragment	closed	Close to F84*	canonical	5GPI 5H4I	1.6 1.2	[17]
Zika G_4SG_4 linker	unobservable after T68*	Solvent-exposed	canonical	5GXJ	2.6	[18]
Zika G_4SG_4 linker	unobservable after A65*	Solvent-exposed	canonical	5T1V	3.1	[19]
Zika G_4SG_4 linker	unobservable after G69*	Solvent-exposed	canonical	5TFN	3	[20]

Zika + inhibitor	closed	Close to F84*	W50 flipped	5H6V	2.4	[21]
Zika + inhibitor	closed	Close to F84*	W50 in two conformations	5YOF	1.5	[22]
Zika + inhibitor	closed	Close to F84*	canonical	5YOD 5ZMQ 5ZMS	1.9 2.0 1.8	[23]
Zika + inhibitor	closed	Close to F84*	W50 in two conformations	5ZOB	2.0	[23]
Zika + inhibitor	closed	Close to F84*	W50 in two conformations	6JPW	1.95	[24]
Zika + inhibitor	closed	Close to F84*	canonical	7DOC	1.9	[25]
Zika + inhibitor	closed	Close to F84*	canonical	6KK2 6KK5	2.0 2.0	[26]
Zika + inhibitor	closed	Close to F84*	W50 flipped	6KK3 6KK4 6KK6 6KPQ 6Y3B	2.05 1.74 1.74 2.6 1.6	[26]
Zika + inhibitor	closed	Close to F84*	canonical	6L50 6L4Z	1.95 1.9	[27]
Zika + inhibitor, G_4SG_4 linker	unobservable after T68*	Solvent-exposed	canonical	6UM3	2.5 Å	[20]
Zika + inhibitor, G_4SG_4 linker	unobservable after T68*	Solvent-exposed, close to W50 of NS3	W50 flipped	7M1V	1.6 Å	[20]
Zika + inhibitor	closed	Close to F84*	canonical	7O55 7OBV 7PFQ 7PFY 7PFZ 7PG1 7VLI 7O2M 7ZNO	1.95 1.3 1.45 1.4 1.45 1.95 2.4 1.9 1.7	[28]
Zika + inhibitor	closed	Close to F84*	W50 flipped	7OC2 7VLG 7PGC 7VLH	1.7 1.7 1.55 2.6	[28]
Zika + inhibitor	closed	Close to F84*	canonical	7VXX 7VXY	1.9 1.9	[29]
Zika + inhibitor	closed	Close to F84*	canonical	7ZLC 7ZLD	1.75 1.6	[30]
Zika + inhibitor	closed	Close to F84*	W50 in two conformations	7ZMI	2.15	[30]

Zika	closed	Close to F84*	W50 in two conformations	8PN6	1.6	Ni et al., to be published
Zika + inhibitor	closed		canonical	7ZQ1 7ZWK	1.5 2.0	[31]
Zika + inhibitor	closed		W50 in two conformations	7ZTM 8A15	1.45 1.23	[31]
Zika + inhibitor	closed	Close to F84*	W50 flipped	7ZPD 7ZQF 7ZUM 7ZV4 7ZVV 7ZW5 7ZYS 8AQA 8AQB 8AQK	1.4 1.7 1.75 1.7 1.75 1.4 1.26 1.35 1.3 1.2	[31]
Yellow fever G_4SG_4 linker	closed	F118 is close to F84*	canonical	6URV	2.9	[32]
Murray Valley encephalitis G_4SG_4 linker	unobservable after T68*	Solvent-exposed	canonical	2WV9	2.75	[33]

[a] NS2Bc refers to the polypeptide segment forming a β-hairpin, which in the closed conformation contributes to the substrate binding site.

[b] The canonical side chain conformations are defined by the high-resolution structure 2FOM. A star marks sequence numbers in NS2B.

Table S2. Mutations found in G1PylRS variants identified to recognize 2F-Phe, 4F-Phe, and F_5-Phe.

RS Variants	*Randomized Sites*						
G1PylRS-wt	L124	Y125	N165	V167	Y204	A221	W237
2FF08	L	Y	D	S	F	C	W
2FF13	L	Y	A	S	Y	C	T
2FF23	A	Y	T	F	Y	C	N
2FF31	L	V	T	F	Y	A	S
2FF44[a]	L	Y	A	F	Y	R	S
2FF51	L	Y	A	F	Y	A	G
4FF02[a]	G	M	N	F	F	A	Y
4FF08	G	F	N	L	W	S	Y
4FF48	N	F	N	F	F	A	Y
F5F09	V	C	T	G	Y	G	W
F5F29[a]	L	Y	N	L	W	S	W
F5F46	C	R	A	S	Y	C	W
F5F54	L	Y	N	F	Y	N	W

[a] The mutants G1-2FF44 and G1-4FF02 were subsequently used for all applications incorporating 2F-Phe and 4F-Phe, respectively. The mutant G1-F5F29 was used for the incorporation of 3F-Phe, 2,6diF-Phe, and 3,5diF-Phe.

Table S3. Protein yields of NS2B-NS3 proteases with fluorinated phenylalanines in position 116.

Protein	Yield in mg purified protein per liter cell culture
gDENV with 2F-Phe	43
gDENV with 3F-Phe	12
gDENV with 4F-Phe	32
gDENV with 3,5diF-Phe	17
gDENV with 2,6diF-Phe	17
eDENV with 2F-Phe	21
eDENV with 3F-Phe	9
eDENV with 4F-Phe	12
eDENV with 3,5diF-Phe	12
eDENV with 2,6diF-Phe	22
gZika with 2F-Phe	16
gZika with 3F-Phe	7
gZika with 4F-Phe	13
gZika with 3,5diF-Phe	15
gZika with 2,6diF-Phe	6
eZika with 2F-Phe	11
eZika with 3F-Phe	5
eZika with 4F-Phe	6
eZika with 3,5diF-Phe	5
eZika with 2,6diF-Phe	7

Table S4. DNA and corresponding amino acid sequences of the proteins used in the current study.

Protein	DNA sequence	Amino acid sequence[a]
G1-2FF44	ATGGTGGTGAAATTTACCGATAGCCAGATTCAGCATCTGATGG AATATGGTGATAATGATTGGAGCGAAGCCGAATTTGAAGATGC AGCAGCACGTGATAAAGAATTTAGCAGCCAGTTTAGCAAACTG AAAAGCGCCAATGATAAAGGCCTGAAAGATGTTATTGCAAATC CGCGTAATGATCTGACCGATCTGGAAAACAAAATTCGCGAAAA ACTGGCAGCCCGTGGTTTTATTGAAGTTCATACCCCGATTTTT GTGAGCAAAAGCGCACTGGCAAAAATGACCATTACCGAAGATC ATCCGCTGTTCAAACAGGTGTTTTGGATTGATGATAAACGTGC ACTGCGTCCGATGCATGCAATGAATCTGTATAAAGTTATGCGT GAACTGCGCGATCATACCAAAGGTCCGGTTAAAATCTTTGAAA TTGGTAGCTGCTTTCGCAAAGAAAGCAAAAGCAGTACCCATCT GGAAGAATTTACCATGCTGGCCCTGTTCGAAATGGGTCCTGAT GGTGATCCGATGGAACATCTGAAAATGTATATTGGCGATATCA TGGATGCCGTTGGTGTTGAATATACCACCAGTCGTGAAGAATC AGATGTTTATGTTGAAACCCTGGACGTGGAAATTAATGGCACC GAAGTTGCAAGCGGTCGGGTTGGTCCGCATAAACTGGATCCGG CACATGATGTGCATGAACCGTCTGCAGGTATTGGTTTTGGTCT GGAACGTCTGCTGATGCTGAAAAATGGTAAAAGCAATGCACGC AAAACCGGCAAAAGTATTACCTATCTGAATGGCTACAAACTGG ATTAA	MVVKFTDSQIQHLMEYGDNDWSEAE FEDAAARDKEFSSQFSKLKSANDKG LKDVIANPRNDLTDLENKIREKLAA RGFIEVHTPIFVSKSALAKMTITED HPLFKQVFWIDDKRALRPMHAMNLY KVMRELRDHTKGPVKIFEIGSCFRK ESKSSTHLEEFTMLALFEMGPDGDP MEHLKMYIGDIMDAVGVEYTTSREE SDVYVETLDVEINGTEVASGRVGPH KLDPAHDVHEPSAGIGFGLERLLML KNGKSNARKTGKSITYLNGYKLD
G1-4FF02	ATGGTGGTGAAATTTACCGATAGCCAGATTCAGCATCTGATGG AATATGGTGATAATGATTGGAGCGAAGCCGAATTTGAAGATGC AGCAGCACGTGATAAAGAATTTAGCAGCCAGTTTAGCAAACTG AAAAGCGCCAATGATAAAGGCCTGAAAGATGTTATTGCAAATC CGCGTAATGATCTGACCGATCTGGAAAACAAAATTCGCGAAAA ACTGGCAGCCCGTGGTTTTATTGAAGTTCATACCCCGATTTTT GTGAGCAAAAGCGCACTGGCAAAAATGACCATTACCGAAGATC ATCCGCTGTTCAAACAGGTGTTTTGGATTGATGATAAACGTGC ACTGCGTCCGATGCATGCAATGAATGGTATGAAAGTTATGCGT GAACTGCGCGATCATACCAAAGGTCCGGTTAAAATCTTTGAAA TTGGTAGCTGCTTTCGCAAAGAAAGCAAAAGCAGTACCCATCT GGAAGAATTTACCATGCTGAACCTGTTCGAAATGGGTCCTGAT GGTGATCCGATGGAACATCTGAAAATGTATATTGGCGATATCA TGGATGCCGTTGGTGTTGAATATACCACCAGTCGTGAAGAATC AGATGTTTTTGTTGAAACCCTGGACGTGGAAATTAATGGCACC GAAGTTGCAAGCGGTGCTGTTGGTCCGCATAAACTGGATCCGG CACATGATGTGCATGAACCGTATGCAGGTATTGGTTTTGGTCT GGAACGTCTGCTGATGCTGAAAAATGGTAAAAGCAATGCACGC AAAACCGGCAAAAGTATTACCTATCTGAATGGCTACAAACTGG ATTAA	MVVKFTDSQIQHLMEYGDNDWSEAE FEDAAARDKEFSSQFSKLKSANDKG LKDVIANPRNDLTDLENKIREKLAA RGFIEVHTPIFVSKSALAKMTITED HPLFKQVFWIDDKRALRPMHAMNGM KVMRELRDHTKGPVKIFEIGSCFRK ESKSSTHLEEFTMLNLFEMGPDGDP MEHLKMYIGDIMDAVGVEYTTSREE SDVFVETLDVEINGTEVASGAVGPH KLDPAHDVHEPYAGIGFGLERLLML KNGKSNARKTGKSITYLNGYKLD
G1-F5F29	ATGGTGGTGAAATTTACCGATAGCCAGATTCAGCATCTGATGG AATATGGTGATAATGATTGGAGCGAAGCCGAATTTGAAGATGC AGCAGCACGTGATAAAGAATTTAGCAGCCAGTTTAGCAAACTG AAAAGCGCCAATGATAAAGGCCTGAAAGATGTTATTGCAAATC CGCGTAATGATCTGACCGATCTGGAAAACAAAATTCGCGAAAA ACTGGCAGCCCGTGGTTTTATTGAAGTTCATACCCCGATTTTT GTGAGCAAAAGCGCACTGGCAAAAATGACCATTACCGAAGATC ATCCGCTGTTCAAACAGGTGTTTTGGATTGATGATAAACGTGC ACTGCGTCCGATGCATGCAATGAATCTGTATAAAGTTATGCGT GAACTGCGCGATCATACCAAAGGTCCGGTTAAAATCTTTGAAA TTGGTAGCTGCTTTCGCAAAGAAAGCAAAAGCAGTACCCATCT GGAAGAATTTACCATGCTGAACCTGTTAGAAATGGGTCCTGAT GGTGATCCGATGGAACATCTGAAAATGTATATTGGCGATATCA TGGATGCCGTTGGTGTTGAATATACCACCAGTCGTGAAGAATC AGATGTTTGGGTTGAAACCCTGGACGTGGAAATTAATGGCACC	MVVKFTDSQIQHLMEYGDNDWSEAE FEDAAARDKEFSSQFSKLKSANDKG LKDVIANPRNDLTDLENKIREKLAA RGFIEVHTPIFVSKSALAKMTITED HPLFKQVFWIDDKRALRPMHAMNLY KVMRELRDHTKGPVKIFEIGSCFRK ESKSSTHLEEFTMLNLLEMGPDGDP MEHLKMYIGDIMDAVGVEYTTSREE SDVWVETLDVEINGTEVASGSVGPH KLDPAHDVHEPWAGIGFGLERLLML KNGKSNARKTGKSITYLNGYKLD

	GAAGTTGCAAGCGGTAGTGTTGGTCCGCATAAACTGGATCCGG CACATGATGTGCATGAACCGTGGGCAGGTATTGGTTTTGGTCT GGAACGTCTGCTGATGCTGAAAAATGGTAAAAGCAATGCACGC AAAACCGGCAAAAGTATTACCTATCTGAATGGCTACAAACTGG ATTAA	
gDENV-F116TAG	ATGGCAGATCTGGAACTGGAACGTGCAGCAGATGTTCGTTGGG AAGAACAGGCAGAAATTAGCGGTAGCAGCCCGATTCTGAGCAT TACCATTAGCGAAGATGGTAGCATGAGCATCAAAAACGAAGAA GAAGAACAGACCCTGGGTGGAGGTGGCTCAGGAGGGGGTGGGG CCGGTGTTCTGTGGGATGTTCCGAGTCCGCCTCCGGTTGGTAA AGCAGAACTGGAAGATGGTGCCTATCGTATTAAACAGAAAGGT ATTCTGGGTTACAGCCAGATTGGTGCGGGTGTTTATAAAGAAG GCACCTTTCATACCATGTGGCATGTTACCCGTGGTGCAGTTCT GATGCATAAAGGTAAACGTATTGAACCGAGCTGGGCAGATGTT AAAAAGGATCTGATTAGCTATGGTGGTGGTTAGAAACTGGAAG GTGAATGGAAAGAAGGTGAAGAAGTTCAGGTTCTGGCCCTGGA ACCGGGTAAAAATCCGCGTGCAGTTCAGACCAAACCGGGTCTG TAGAAAACCAATACCGGCACCATTGGTGCAGTGAGCCTGGATT TTAGTCCGGGTACAAGCGGTAGCCCGATTGTTGATAAAAAGGG TAAAGTTGTTGGCCTGTATGGTAATGGTGTTGTGACCCGTAGC GGTGCATATGTTAGCGCAATTGCAAATACCGAAAAACATCACC ACCATCATCACTAA	MADLELERAADVRWEEQAEISGSSP ILSITISEDGSMSIKNEEEEQTLGG GGSGGGGAGVLWDVPSPPPVGKAEL EDGAYRIKQKGILGYSQIGAGVYKE GTFHTMWHVTRGAVLMHKGKRIEPS WADVKKDLISYGGGWKLEGEWKEGE EVQVLALEPGKNPRAVQTKPGLXKT NTGTIGAVSLDFSPGTSGSPIVDKK GKVVGLYGNGVVTRSGAYVSAIANT EKHHHHHH
eDENV-F116TAG	ATGGCAGATCTGGAACTGGAACGTGCAGCAGATGTTCGTTGGG AAGAACAGGCAGAAATTAGCGGTAGCAGCCCGATTCTGAGCAT TACCATTAGCGAAGATGGTAGCATGAGCATCAAAAACGAAGAA GAAGAACAGACCCTGGAAGTTAAAAAGCAGCGTGCCGGTGTTC TGTGGGATGTTCCGAGTCCGCCTCCGGTTGGTAAAGCAGAACT GGAAGATGGTGCCTATCGTATTAAACAGAAAGGTATTCTGGGT TACAGCCAGATTGGTGCGGGTGTTTATAAAGAAGGCACCTTTC ATACCATGTGGCATGTTACCCGTGGTGCAGTTCTGATGCATAA AGGTAAACGTATTGAACCGAGCTGGGCAGATGTTAAAAAGGAT CTGATTAGCTATGGTGGTGGTTAGAAACTGGAAGGTGAATGGA AAGAAGGTGAAGAAGTTCAGGTTCTGGCCCTGGAACCGGGTAA AAATCCGCGTGCAGTTCAGACCAAACCGGGTCTGTAGAAAACC AATACCGGCACCATTGGTGCAGTGAGCCTGGATTTTAGTCCGG GTACAAGCGGTAGCCCGATTGTTGATAAAAAGGGTAAAGTTGT TGGCCTGTATGGTAATGGTGTTGTGACCCGTAGCGGTGCATAT GTTAGCGCAATTGCAAATACCGAAAAACATCACCACCATCATC ACTAA	MADLELERAADVRWEEQAEISGSSP ILSITISEDGSMSIKNEEEEQTLEV KKQRAGVLWDVPSPPPVGKAELEDG AYRIKQKGILGYSQIGAGVYKEGTF HTMWHVTRGAVLMHKGKRIEPSWAD VKKDLISYGGGWKLEGEWKEGEEVQ VLALEPGKNPRAVQTKPGLXKTNTG TIGAVSLDFSPGTSGSPIVDKKGKV VGLYGNGVVTRSGAYVSAIANTEKH HHHHH

gZika-F116TAG	ATGTCGGTAGATATGTACATCGAACGCGCAGGAGACATTACCT GGGAGAAGGACGCTGAAGTCACTGGCAATTCACCCCGTTTAGA CGTTGCTCTGGATGAATCGGGTGACTTTAGTCTTGTAGAAGAC GATGGTCCTCCAATGGCGGGTGGAGGTGGCTCAGGAGGGGGTG GGTCTGGTGCATTATGGGACGTACCAGCACCTAAGGAAGTAAA AAACGGCGAAACTACAGACGGTGTGTATCGCGTGATGACACGT GGTTTGTTGGGTAGCACACAAGTGGGCGTCGGAGTAATGCAAG AGGGCGTATTTCACACGATGTGGCATGTAACCAAAGGTTCTGC GTTACGCTCCGGCGAAGGACGCCTTGACCCGTATTGGGGCGAC GTCAAGCAAGATTTGGTTAGCTATAGCGGACCTTAGAAACTGG ACGCGGCATGGGACGGCCACTCGGAGGTACAGTTATTGGCCGT TCCGCCAGGAGAGCGTGCTCGTAACATTCAAACGCTGCCTGGT ATTTAGAAAACTAAGGATGGTGATATCGGGGCGGTGGCCTTAG ATTATCCGGCAGGAACATCCGGGTCTCCCATTTTAGACAAGAG TGGTCGCGTGATTGGGTTATATGGGAATGGGGTTGTAATCAAG AATGGATCTTACGTCTCGGCGATCACACAGGGTCGCCGCCATC ACCACCATCATCACTAA	MSVDMYIERAGDITWEKDAEVTGNS PRLDVALDESGDFSLVEDDGPPMAG GGGSGGGGSGALWDVPAPKEVKNGE TTDGVYRVMTRGLLGSTQVGVGVMQ EGVFHTMWHVTKGSALRSGEGRLDP YWGDVKQDLVSYSGPWKLDAAWDGH SEVQLLAVPPGERARNIQTLPGIXK TKDGDIGAVALDYPAGTSGSPILDK SGRVIGLYGNGVVIKNGSYVSAITQ GRRHHHHHH
eZika-F116TAG	ATGTCGGTAGATATGTACATCGAACGCGCAGGAGACATTACCT GGGAGAAGGACGCTGAAGTCACTGGCAATTCACCCCGTTTAGA CGTTGCTCTGGATGAATCGGGTGACTTTAGTCTTGTAGAAGAC GATGGTCCTCCAATGGCGGTTAAAACCGGTCGTAAATCTGGTG CATTATGGGACGTACCAGCACCTAAGGAAGTAAAAAACGGCGA AACTACAGACGGTGTGTATCGCGTGATGACACGTGGTTTGTTG GGTAGCACACAAGTGGGCGTCGGAGTAATGCAAGAGGGCGTAT TTCACACGATGTGGCATGTAACCAAAGGTTCTGCGTTACGCTC CGGCGAAGGACGCCTTGACCCGTATTGGGGCGACGTCAAGCAA GATTTGGTTAGCTATAGCGGACCTTAGAAACTGGACGCGGCAT GGGACGGCCACTCGGAGGTACAGTTATTGGCCGTTCCGCCAGG AGAGCGTGCTCGTAACATTCAAACGCTGCCTGGTATTTAGAAA ACTAAGGATGGTGATATCGGGGCGGTGGCCTTAGATTATCCGG CAGGAACATCCGGGTCTCCCATTTTAGACAAGAGTGGTCGCGT GATTGGGTTATATGGGAATGGGGTTGTAATCAAGAATGGATCT TACGTCTCGGCGATCACACAGGGTCGCCGCCATCACCACCATC ATCACTAA	MSVDMYIERAGDITWEKDAEVTGNS PRLDVALDESGDFSLVEDDGPPMAV KTGRKSGALWDVPAPKEVKNGETTD GVYRVMTRGLLGSTQVGVGVMQEGV FHTMWHVTKGSALRSGEGRLDPYWG DVKQDLVSYSGPWKLDAAWDGHSEV QLLAVPPGERARNIQTLPGIXKTKD GDIGAVALDYPAGTSGSPILDKSGR VIGLYGNGVVIKNGSYVSAITQGRR HHHHHH

[a] X indicates the position of fluoro-phenylalanines. The Gly_4-Ser-Gly_4 linkers and peptide segments for autodigestion in the NS2B-NS3 protease constructs are colored in blue.

References

(1) Abdelkader, E. H.; Qianzhu, H.; George, J.; Frkic, R. L.; Jackson, C. J.; Nitsche, C.; Otting, G.; Huber, T. Genetic encoding of cyanopyridylalanine for in-cell protein macrocyclization by the nitrile–aminothiol click reaction. *Angew. Chem. Int. Ed.* **2022**, *61*, e202114154.

(2) Qianzhu, H.; Abdelkader, E. H.; Herath, I. D.; Otting, G.; Huber, T. Site-specific incorporation of 7-fluoro-L-tryptophan into proteins by genetic encoding to monitor ligand binding by ^{19}F NMR spectroscopy. *ACS Sens.* **2022**, *7*, 44–49.

(3) Mukai, T.; Hoshi, H.; Ohtake, K.; Takahashi, M.; Yamaguchi, A.; Hayashi, A.; Yokoyama, S.; Sakamoto, K. Highly reproductive *Escherichia coli* cells with no specific assignment to the UAG codon. *Sci. Rep.* **2015**, *5*, 9699.

(4) Phoo, W. W.; El Sahili, A.; Zhang, Z.; Chen, M. W.; Liew, C. W.; Lescar, J.; Vasudevan, S. G.; Luo, D. Crystal structures of full length DENV4 NS2B-NS3 reveal the dynamic interaction between NS2B and NS3. *Antiviral Res.* **2020**, *182*, 104900.

(5) Quek, J.-P.; Ser, Z.; Chew, B. L. A.; Li, X.; Wang, L.; Sobota, R. M.; Luo, D.; Phoo, W. W. Dynamic interactions of post cleaved NS2B cofactor and NS3 protease identified by integrative structural approaches. *Viruses* **2022**, *14*, 1440.

(6) Noble, C. G.; Seh, C. C.; Chao, A. T.; Shi, P. Y. Ligand-bound structures of the dengue virus protease reveal the active conformation. *J. Virol.* **2012**, *86*, 438–446.

(7) Erbel, P.; Schiering, N.; D'Arcy, A.; Renatus, M.; Kroemer, M.; Lim, S. P.; Yin, Z.; Keller, T. H.; Vasudevan, S. G.; Hommel, U. Structural basis for the activation of flaviviral NS3 proteases from dengue and West Nile virus. *Nat. Struct. Mol. Biol.* **2006**, *13*, 372–373.

(8) Yildiz, M.; Ghosh, S.; Bell, J. A.; Sherman, W.; Hardy, J. A. Allosteric inhibition of the NS2B-NS3 protease from dengue virus. *ACS Chem. Biol.* **2013**, *8*, 2744–2752.

(9) Yao, Y.; Huo, T.; Lin, Y.-L.; Nie, S.; Wu, F.; Hua, Y.; Wu, J.; Kneubehl, A. R.; Vogt, M. B.; Rico-Hesse, R.; Song, Y. Discovery, X-ray crystallography and antiviral activity of allosteric inhibitors of flavivirus NS2B-NS3 protease. *J. Am. Chem. Soc.* **2019**, *141*, 6832–6836.

(10) Chandramouli, S.; Joseph Jeremiah, S.; Daudenarde, S.; Gatchalian, J.; Cornillez-Ty, C.; Kuhn, P. Serotype-specific structural differences in the protease-cofactor complexes of the dengue virus family. *J. Virol.* **2010**, *84*, 3059–3067.

(11) Aleshin, A. E.; Shiryaev, S. A.; Strongin, A. Y.; Liddington, R. C. Structural evidence for regulation and specificity of flaviviral proteases and evolution of the *Flaviviridae* fold. *Protein Sci.* **2007**, *16*, 795–806.

(12) Robin, G.; Chappell, K.; Stoermer, M. J.; Hu, S.-H.; Young, P. R.; Fairlie, D. P.; Martin, J. L. Structure of West Nile virus NS3 protease: ligand stabilization of the catalytic conformation. *J. Mol. Biol.* **2009**, *385*, 1568–1577.

(13) Hammamy, M. Z.; Haase, C.; Hammami, M.; Hilgenfeld, R.; Steinmetzer, T. Development and characterization of new peptidomimetic inhibitors of the West Nile virus NS2B–NS3 protease. *ChemMedChem* **2013**, *8*, 231–241.

(14) Nitsche, C.; Zhang, L.; Weigel, L. F.; Schilz, J.; Graf, D.; Bartenschlager, R.; Hilgenfeld, R.; Klein, C. D. Peptide–boronic acid inhibitors of flaviviral proteases: medicinal chemistry and structural biology. *J. Med. Chem.* **2017**, *60*, 511–516.

(15) Lei, J.; Hansen, G.; Nitsche, C.; Klein, C. D.; Zhang, L.; Hilgenfeld, R. Crystal structure of Zika virus NS2B-NS3 protease in complex with a boronate inhibitor. *Science* **2016**, *353*, 503–505.

(16) Phoo, W. W.; Li, Y.; Zhang, Z.; Lee, M. Y.; Loh, Y. R.; Tan, Y. B.; Ng, E. Y.; Lescar, J.; Kang, C.; Luo, D. Structure of the NS2B-NS3 protease from Zika virus after self-cleavage. *Nat. Commun.* **2016**, *7*, 13410.

(17) Zhang, Z.; Li, Y.; Loh, Y. R.; Phoo, W. W.; Hung, A. W.; Kang, C.; Luo, D. Crystal structure of unlinked NS2B-NS3 protease from Zika virus. *Science* **2016**, *354*, 1597–1600.

(18) Chen, X.; Yang, K.; Wu, C.; Chen, C.; Hu, C.; Buzovetsky, O.; Wang, Z.; Ji, X.; Xiong, Y.; Yang, H. Mechanisms of activation and inhibition of Zika virus NS2B-NS3 protease. *Cell Res.* **2016**, *26*, 1260–1263.

(19) Lee, H.; Ren, J.; Nocadello, S.; Rice, A. J.; Ojeda, I.; Light, S.; Minasov, G.; Vargas, J.; Nagarathnam, D.; Anderson, W. F.; Johnson, M. E. Identification of novel small molecule inhibitors against NS2B/NS3 serine protease from Zika virus. *Antiviral Res.* **2017**, *139*, 49–58.

(20) Shiryaev, S. A.; Cieplak, P.; Cheltsov, A.; Liddington, R. C.; Terskikh, A. V. Dual function of Zika virus NS2B-NS3 protease. *PLoS Path.* **2023**, *19*, e1011795.

(21) Li, Y.; Zhang, Z.; Phoo, W. W.; Loh, Y. R.; Wang, W.; Liu, S.; Chen, M. W.; Hung, A. W.; Keller, T. H.; Luo, D.; Kang, C. Structural dynamics of Zika virus NS2B-NS3 protease binding to dipeptide inhibitors. *Structure* **2017**, *25*, 1242–1250.e3.

(22) Li, Y.; Zhang, Z.; Phoo, W. W.; Loh, Y. R.; Li, R.; Yang, H. Y.; Jansson, A. E.; Hill, J.; Keller, T. H.; Nacro, K.; Luo, D.; Kang, C. Structural insights into the inhibition of Zika virus NS2B-NS3 protease by a small-molecule inhibitor. *Structure* **2018**, *26*, 555–564.e3.

(23) Phoo, W. W.; Zhang, Z.; Wirawan, M.; Chew, E. J. C.; Chew, A. B. L.; Kouretova, J.; Steinmetzer, T.; Luo, D. Structures of Zika virus NS2B-NS3 protease in complex with peptidomimetic inhibitors. *Antiviral Res.* **2018**, *160*, 17–24.

(24) Nitsche, C.; Onagi, H.; Quek, J.-P.; Otting, G.; Luo, D.; Huber, T. Biocompatible macrocyclization between cysteine and 2-cyanopyridine generates stable peptide inhibitors. *Org. Lett.* **2019**, *21*, 4709–4712.
(25) Patil, N. A.; Quek, J.-P.; Schroeder, B.; Morewood, R.; Rademann, J.; Luo, D.; Nitsche, C. 2-Cyanoisonicotinamide conjugation: a facile approach to generate potent peptide inhibitors of the Zika virus protease. *ACS Med. Chem. Lett.* **2021**, *12*, 732–737.
(26) Braun, N. J.; Quek, J. P.; Huber, S.; Kouretova, J.; Rogge, D.; Lang-Henkel, H.; Cheong, E. Z. K.; Chew, B. L. A.; Heine, A.; Luo, D.; Steinmetzer, T. Structure-based macrocyclization of substrate analogue NS2B-NS3 protease inhibitors of Zika, West Nile and dengue viruses. *ChemMedChem* **2020**, *15*, 1439–1452.
(27) Quek, J. P.; Liu, S.; Zhang, Z.; Li, Y.; Ng, E. Y.; Loh, Y. R.; Hung, A. W.; Luo, D.; Kang, C. Identification and structural characterization of small molecule fragments targeting Zika virus NS2B-NS3 protease. *Antiviral Res.* **2020**, *175*, 104707.
(28) Huber, S.; Braun, N. J.; Schmacke, L. C.; Quek, J. P.; Murra, R.; Bender, D.; Hildt, E.; Luo, D.; Heine, A.; Steinmetzer, T. Structure-based optimization and characterization of macrocyclic Zika virus NS2B-NS3 protease inhibitors. *J. Med. Chem.* **2022**, *65*, 6555–6572.
(29) Xiong, Y.; Cheng, F.; Zhang, J.; Su, H.; Hu, H.; Zou, Y.; Li, M.; Xu, Y. Structure-based design of a novel inhibitor of the ZIKA virus NS2B/NS3 protease. *Bioorg. Chem.* **2022**, *128*, 106109.
(30) Hammerschmidt, S. J.; Huber, S.; Braun, N. J.; Lander, M.; Steinmetzer, T.; Kersten, C. Thermodynamic characterization of a macrocyclic Zika virus NS2B/NS3 protease inhibitor and its acyclic analogs. *Arch. Pharm.* **2023**, *356*, 2200518.
(31) Huber, S.; Braun, N. J.; Schmacke, L. C.; Murra, R.; Bender, D.; Hildt, E.; Heine, A.; Steinmetzer, T. Synthesis and structural characterization of new macrocyclic inhibitors of the Zika virus NS2B–NS3 protease. *Arch. Pharm.* **2024**, *357*, 2400250.
(32) Noske, G. D.; Gawriljuk, V. O.; Fernandes, R. S.; Furtado, N. D.; Bonaldo, M. C.; Oliva, G.; Godoy, A. S. Structural characterization and polymorphism analysis of the NS2B-NS3 protease from the 2017 Brazilian circulating strain of Yellow Fever virus. *Biochim. Biophys. Acta - Gen. Subj.* **2020**, *1864*, 129521.
(33) Assenberg, R.; Mastrangelo, E.; Walter, T. S.; Verma, A.; Milani, M.; Owens, R. J.; Stuart, D. I.; Grimes, J. M.; Mancini, E. J. Crystal structure of a novel conformational state of the flavivirus NS3 protein: Implications for polyprotein processing and viral replication. *J. Virol.* **2009**, *83*, 12895–12906.