Supporting Information

De Novo Discovery of Nonstandard Macrocyclic Peptides as Noncompetitive Inhibitors of the Zika Virus NS2B-NS3 Protease

Christoph Nitsche, Toby Passioura, Paul Varava, Mithun C. Mahawaththa, Mila M. Leuthold, Christian D. Klein, Hiroaki Suga, and Gottfried Otting

Preparation of linked Zika virus NS2B-NS3 protease (gZiPro). The construct has been described previously.[1, 2] Briefly, it includes 48 hydrophilic core residues of NS2B followed by a GGGGSGGGG linker, the 170 N-terminal residues of the NS3 protease domain and a C-terminal His_6 tag. The mutations C80S and C143S were introduced to avoid dimerization by oxidation. The T7 expression vector pETMCSI[3] was transformed into *E. coli* BL21(DE3) cells, which were grown in LB medium at 37 °C until an OD_{600} value of 0.6 was reached. Subsequently, overexpression was induced with IPTG (1 mM), and the cells were incubated at room temperature overnight. The cells were pelleted by centrifuging at 5,000 *g* for 10 minutes and lysed by passing through a French Press (SLM Aminco) at 830 bars. The cell lysate was centrifuged for 1 h at 34,000 *g* and the supernatant was loaded onto a 5 mL Co-NTA column (GE Healthcare Life Science) pre-equilibrated with buffer A (50 mM Tris-HCl, pH 7.5, 300 mM NaCl, 10% (v/v) glycerol). The protein was eluted with a linear gradient of buffer A supplemented with 300 mM imidazole, fractions were analyzed by 12% SDS-PAGE, and the buffer was exchanged to 50 mM phosphate, pH 7.0, and 50 mM NaCl.

Preparation of unlinked Zika virus NS2B-NS3 protease (bZiPro). The construct was obtained from Addgene (plasmid #86846) and expressed and purified exactly as described previously.[4]

Synthesis and screening of peptide library. Library construction and screening were performed essentially as described previously (Figure S1).[5-11] Briefly, DNA templates for the construction of mRNA libraries were synthesized by primer extension and PCR using the primers shown in Table S2. Cognate mRNAs were generated by T7 polymerase-mediated transcription and covalently linked to PEG-puromycin, before translation in reprogrammed *in vitro* translation reactions without Asp, Glu, Met, and Arg and their cognate aminoacyl-tRNA synthetases, as previously described.[12] Each reaction contained 1.2 µM mRNA-puro, 12.5 µM initiator tRNA ($tRNA^{fMet}$ aminoacylated with ClAc-Tyr), and 25 µM of each elongator tRNA (an engineered Glu tRNA including an anticodon of choice - EnGluxxx, see Table S2) aminoacylated with the specified non-canonical amino acid. In the first round of selection, translation was performed at 150 µL scale to produce approximately $1.2 \cdot 10^{14}$ molecules. After translation, peptide-mRNA conjugates were isolated from the ribosomes by EDTA treatment and reverse transcribed using the CGT3an13.R23 primer and RNase H-reverse transcriptase (Promega). Peptide-mRNA conjugate libraries were panned against 200 nM C-terminally His_6-tagged linked Zika virus NS2B-NS3 protease (gZiPro) immobilized on Dynabeads™ (His-tag isolation and pulldown - Thermo Fisher) at 4 °C for 30 min, the beads were washed three times

with cold PBS-T (10 mM phosphate, pH 7.4, 130 mM NaCl, 0.05% (v/v) Tween-20), and cDNA was recovered by PCR using the primers T7g10m.F46 and CGT3an13.R23 prior to T7 transcription to generate the mRNA for a second round of selection. In the second and later rounds of selection, translation was performed at 5 μL scale and libraries were subjected to 6 rounds of counter selection against uncoated beads prior to panning against linked Zika virus NS2B-NS3 protease (gZiPro). Assessment of recovery following each round was performed by quantitative real time PCR using Sybr Green I dye (Lonza Japan Ltd) on a light cycler nanothermal cycler (Roche). DNA from each round was sequenced using a MiSeq high-throughput sequencer (Illumina) and analyzed using CLC workbench software (Qiagen) as previously described.[9] Following alignment of the 100 most enriched sequences (Figure S2), peptide families were defined as groups of close analogues for which at least 2 sequences were identified, and the most abundant peptide from each was chosen for solid phase synthesis and further characterisation.

Solid phase peptide synthesis. Macrocyclic peptides were synthesized at 25 μmol scale using standard Fmoc solid phase peptide synthesis methodology using HBTU, HOBt, diisopropylethylamine (DIPEA), Rink amide resin, and a Syro I automated synthesizer (Biotage). All amino acids and coupling reagents were purchased from Watanabe Chemical Ltd, except for DIPEA (Wako Japan). For synthesis of *N*-methyl-*O*-methyl-tyrosine residues, *O*-methyl-tyrosine coupling was followed by Fmoc deprotection, nosyl protection of the resulting free amine, and on-resin methylation using dimethyl sulfate as follows: following Fmoc deprotection, 2-nitrobenzenesulfonyl chloride (22 mg, 0.1 mmol) in *N*-methyl-2-pyrrolidone (NMP) and 2,4,6-trimethylpyridine (33 μL, 0.25 mmol) were added to the resin, the reaction was allowed to proceed for 15 min at room temperature, and the resin was washed (five times) with NMP. Methylation was performed twice using 1,8-diazabicyclo(5.4.0)undec-7-ene (11.25 μL, 0.075 mmol) and dimethyl sulfate (22.25 μL, 0.25 mmol) in NMP for 5 minutes, after which the resin was washed (five times) with NMP. Nosyl deprotection was performed twice using 1,8-diazabicyclo(5.4.0)undec-7-ene (18.75 μL, 0.125 mmol) and 2-mercaptoethanol (17.5 μL, 0.25 mmol) in NMP for 5 minutes, after which the resin was washed (five times) with NMP and solid phase peptide synthesis was continued. Identity and purity of all peptides were confirmed by MALDI-TOF MS (Figure S3).

Zika virus NS2B-NS3 protease enzymatic inhibition assay. All measurements were performed in triplicate in 10 mM Tris-HCl, pH 8.5, 20% (v/v) glycerol, and 1 mM CHAPS as described previously.[13, 14] Different concentrations (ranging from 0 to 100 μM) of compounds **1-6** (10 mM DMSO stocks) were pipetted into a 96-well plate (black U-bottom, Greiner, Bio-One). Zika virus NS2B-NS3 protease was added to a final concentration of 1 nM (gZiPro) or 0.5 nM (bZiPro) and the mixture was incubated for 10 min. Subsequently, the enzymatic reaction was initiated by adding the substrate Bz-Nle-Lys-Lys-Arg-AMC (Biosyntan) to four final concentrations (15, 30, 45, 60 μM). The gradual release of fluorescent 7-amino-4-methylcoumarin (AMC) was monitored for 3 min at 460 nm with excitation at 360 nm, using a fluorophotometer (Spectramax M2e plate reader, Molecular Devices). Initial velocities were derived from the linear curves as variation of relative fluorescence intensity per unit time. The relative fluorescence units were converted to the amount of cleaved substrate via a calibration curve as described previously.[14] IC_{50} values were calculated from at least seven different inhibitor concentrations using Prism 7.0 (Graphpad Software). K_i values for a non-competitive inhibition model were calculated using Prism 7.0 (GraphPad Software) using non-linear least-squares fits.

Surface plasmon resonance (SPR) determination of binding kinetics. Binding kinetics of each macrocyclic peptide toward Zika virus NS2B-NS3 protease were determined using a Biacore T200 SPR instrument (GE Healthcare Life Sciences). Target protein was immobilized on a CM5 chip using EDC/NHS chemistry (Amine Coupling Kit, GE Healthcare Life Sciences) to a signal level of ~2000 RU. Binding kinetics were assessed by injection of varying concentrations of compounds **1-6** dissolved in 10 mM HEPES, pH 7.4, 150 mM NaCl, 0.5% (v/v) polysorbate 20, 0.1% (v/v) DMSO with target regeneration using 10 mM glycine buffer, pH 1.5.

Preparation of dengue and West Nile virus NS2B-NS3 proteases. The protease constructs of dengue serotype 2 and West Nile virus were described previously.[15, 16] In both constructs the core sequence of NS2B is covalently ligated to the protease NS3 domain by a GGGGSGGGG linker. Transformation of the pET28a plasmids (Novagen), expression of the His_6-tagged proteins in *E. coli* BL21(DE3) cells, and purification by nickel affinity chromatography were performed following a previously described protocol.[15, 16]

Dengue and West Nile virus NS2B-NS3 protease enzymatic inhibition assays. All measurements were performed in triplicate as described previously.[15, 16] Briefly, continuous assays were performed in 50 mM Tris-HCl, pH 9, 10% (v/v) ethylene glycol, and 0.0016% Brij 58 using a BMG Labtech Fluostar OPTIMA microtiter fluorescence plate reader and black 96 well V-bottom plates (Greiner Bio-One, Germany) with an excitation wavelength of 320 nm and a monitored emission wavelength of 405 nm. Compounds **1**-**6** (10 mM in DMSO) were incubated with the dengue virus protease (100 nM) or West Nile virus protease (150 nM) for 15 min before the enzymatic reaction was initiated by addition of the FRET substrates (final concentration 50 µM) Abz-Nle-Lys-Arg-Arg-Ser-3-(NO_2)Tyr (dengue virus protease) or Abz-Gly-Leu-Lys-Arg-Gly-Gly-3-(NO_2)Tyr (West Nile virus protease). The enzymatic activity was monitored for 15 min and initial velocities were derived from the linear curves as variation of relative fluorescence intensity per unit time. IC_{50} values were calculated from at least seven different inhibitor concentrations using Prism 7.0 (Graphpad Software).

Thrombin and trypsin enzymatic assays. All measurements were performed in triplicate as described previously.[16] Briefly, black 96 well V-bottom plates (Greiner Bio-One), a BMG Labtech Fluostar OPTIMA microtiter fluorescence plate reader operating at an excitation wavelength of 355 nm and an emission wavelength of 460 nm, and a buffer consisting of 50 mM Tris-HCl pH 7.5, 150 mM NaCl, and 0.05% Tween 20 were used. Compounds **1** and **2** were preincubated with thrombin (10 nM) or trypsin (1 nM) for 15 minutes at final concentrations of 25 and 50 µM, respectively. The cleavage reaction was initiated by addition of the Boc-Val-Pro-Arg-AMC substrate (Bachem, Switzerland) at a final concentration of 50 µM. The activities of thrombin and trypsin were determined as fluorescence increase (RFU/s) and monitored for 10 min.

Huh-7 metabolic activity assay. Huh-7 cells were seeded into 96-well cell culture plates at a density of 10^4 cells per well in a final volume of 100 µl. Treatment with compounds **1** and **2** was performed immediately after seeding at a starting concentration of 50 µM and a 2-fold serial dilution. After incubation for 24 h or 48 h at 37 °C, the medium was replaced by fresh DMEM and cell viability was measured by adding 20 µl per well CellTiter-Blue® reagent (Promega), which monitors the metabolic conversion of resazurin (λ_{max} = 605 nm) to resorufin (λ_{max} = 573 nm). Subsequently, cells were incubated for 1 h at 37 °C and the absorbance was measured at 600 and 570 nm. For evaluation the difference in absorbance values (absolute values or relative to the non-treated control) were plotted against the compound concentration.

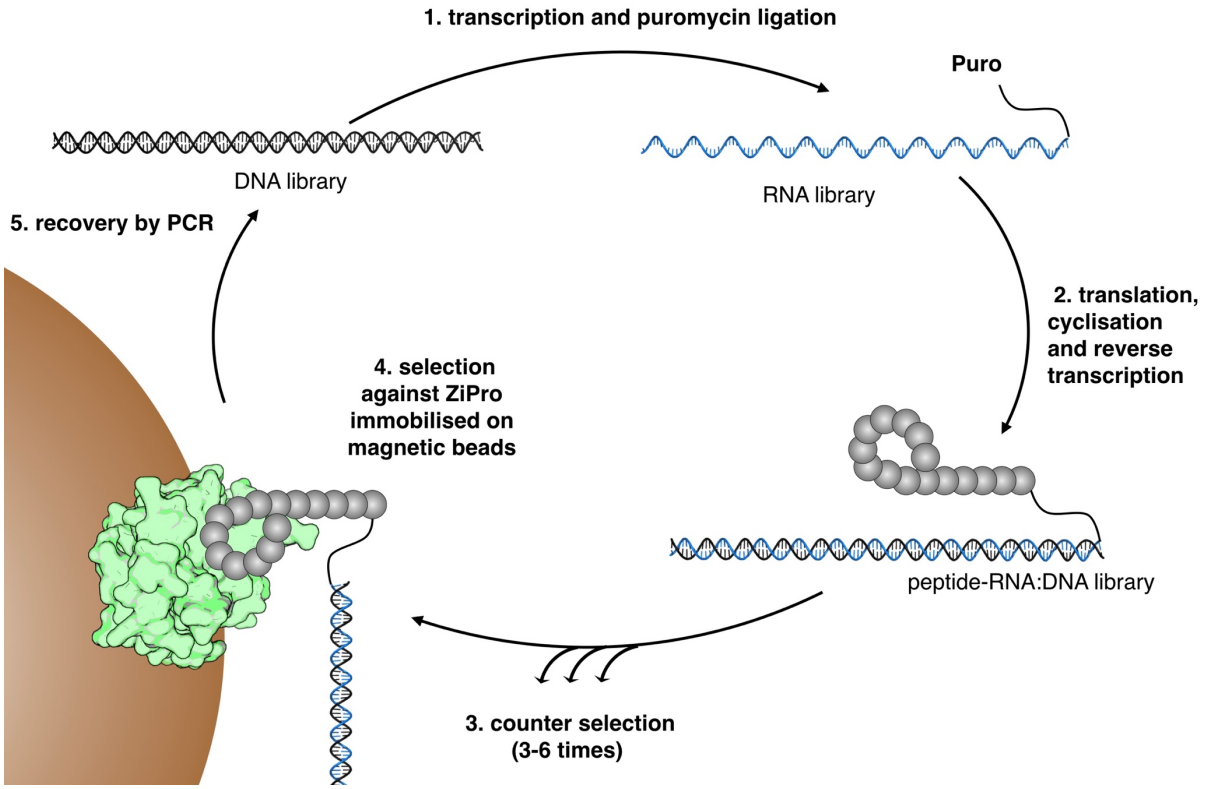
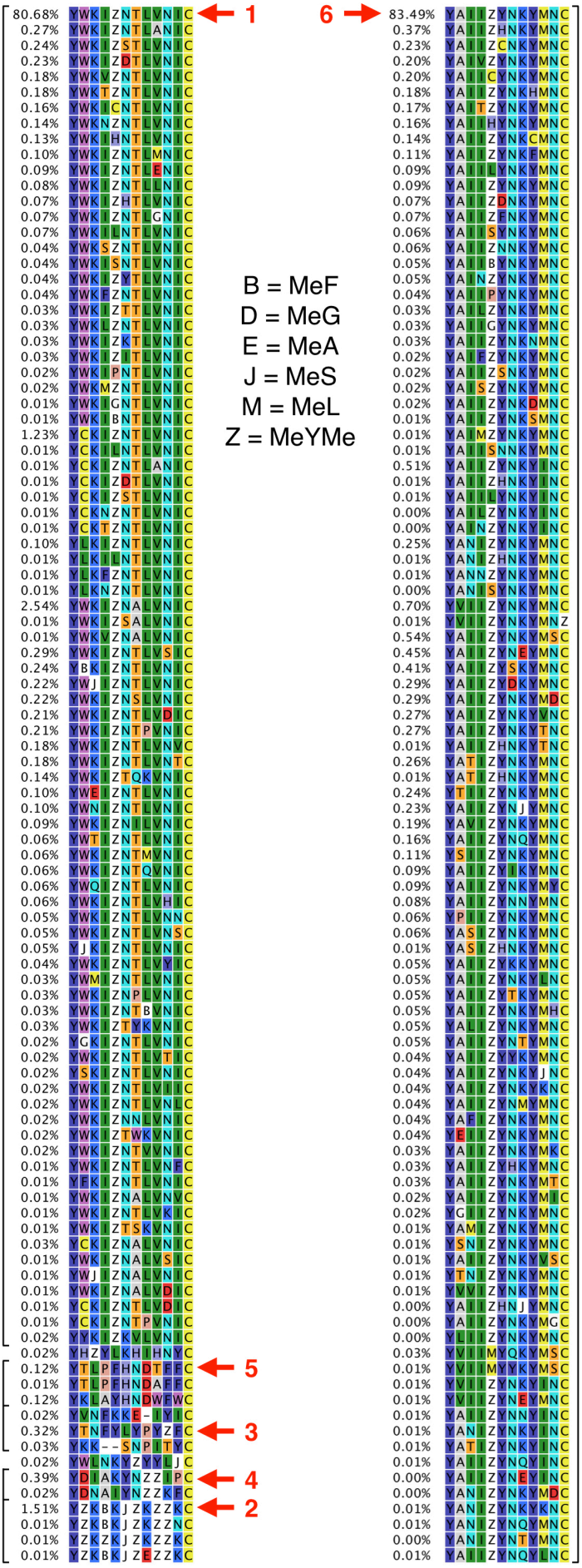


Figure S1. Schematic illustration of RaPID screening process. 1. A DNA library assembled from synthetic oligonucleotides is transcribed into RNA (blue) and ligated to puromycin at the 3' end. **2.** This library is translated in a genetically reprogrammed reaction, resulting in a macrocyclic peptide library in which each peptide is covalently linked to its cognate mRNA. This is then reverse transcribed to generate peptide-RNA:DNA molecules. **3.** Counter selections are used to remove peptides that bind non-specifically to the bead surface. **4.** The library is then panned against Zika virus NS2B-NS3 protease immobilised on magnetic beads. **5.** An enriched DNA library is then recovered by PCR and the process is repeated until increased rates of target binding are observed. Deconvolution of the library can be achieved through sequencing of the final (and intermediate) enriched DNA libraries.

Figure S2. Alignment of peptide sequences identified following selection for Zika virus NS2B-NS3 ligands. The hundred most frequent peptides from each library are shown. Sequences from the ClAc-L-tyrosine initiated library are shown in the left-hand column, and sequences from the ClAc-D-tyrosine initiated library are shown in the right-hand column. Other non-canonical amino acids are represented as indicated in the legend. Percentage values next to each sequence indicate overall prevalence in the enriched library. Black brackets show the breakdown of the identified sequences into 6 families (5 ClAc-L-tyrosine initiated and 1 ClAc-D-tyrosine initiated). Red arrows indicate sequences chosen for synthesis and further characterization, with numbers corresponding to the compounds listed in Table 1.



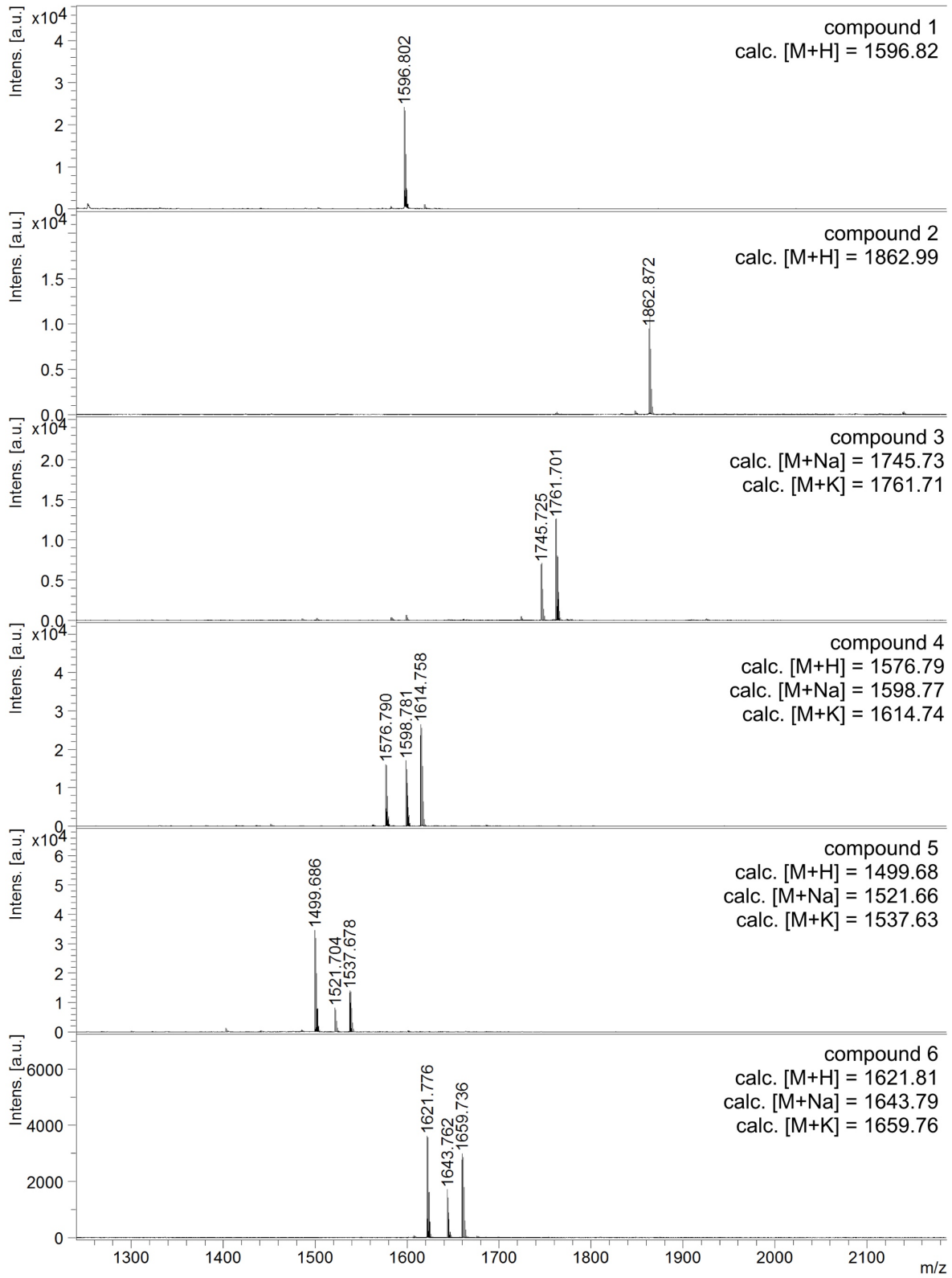


Figure S3. MALDI-TOF MS analysis of synthesized peptides.

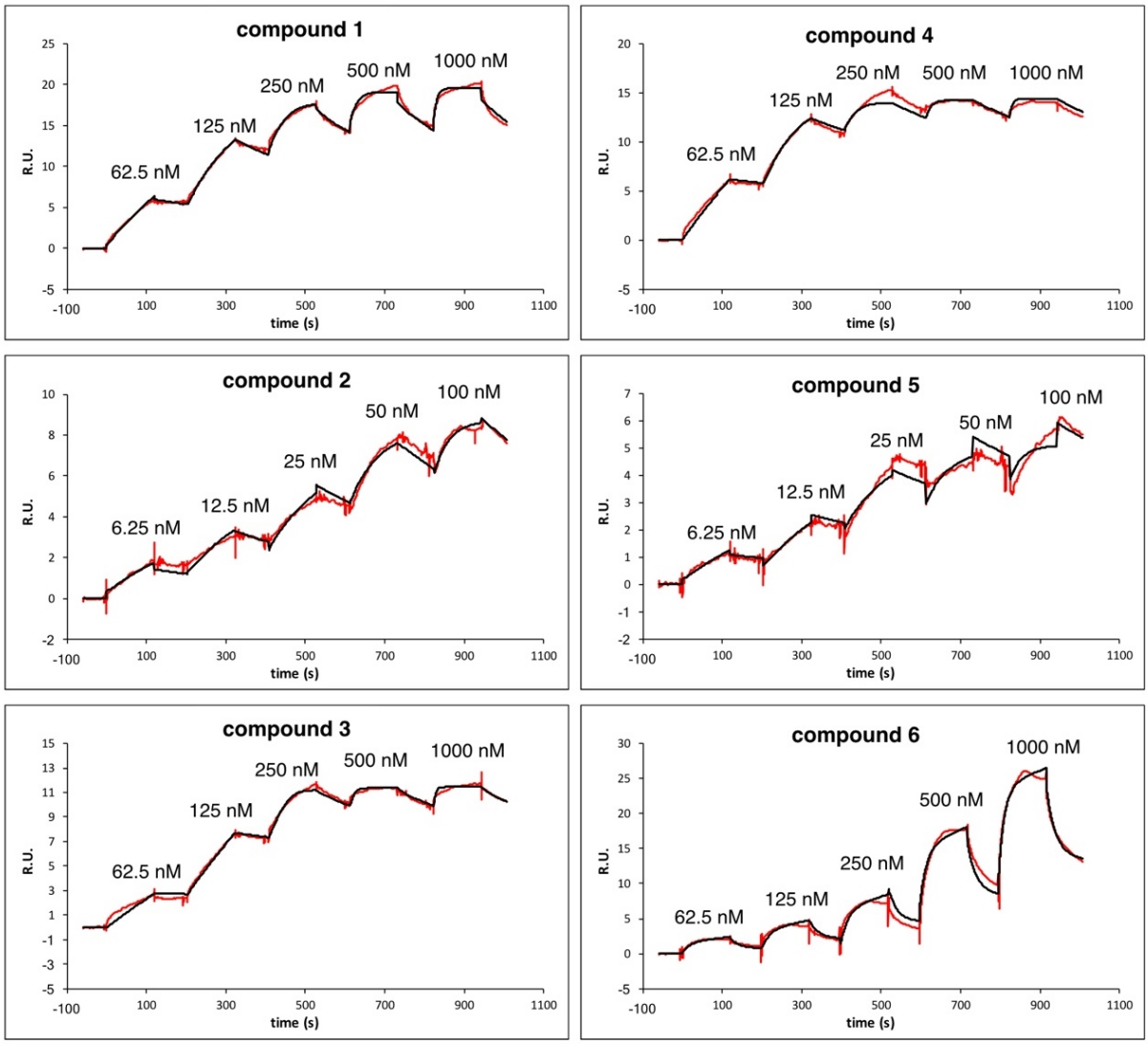


Figure S4. Zika virus NS2B-NS3 protease binding kinetics of compounds 1-6. Single-cycle kinetics experiments were performed using surface immobilized Zika virus NS2B-NS3 protease (gZiPro) with varying concentrations of macrocyclic peptide analyte. Compounds were analysed at the concentrations indicated. Kinetic parameters (see Table S1) were determined by fitting a 1:1 binding model, with the exception of compound **6** which was fitted using a two-state binding model. In each case, the red line indicates the data trace, and the black line indicates the fitted curve.

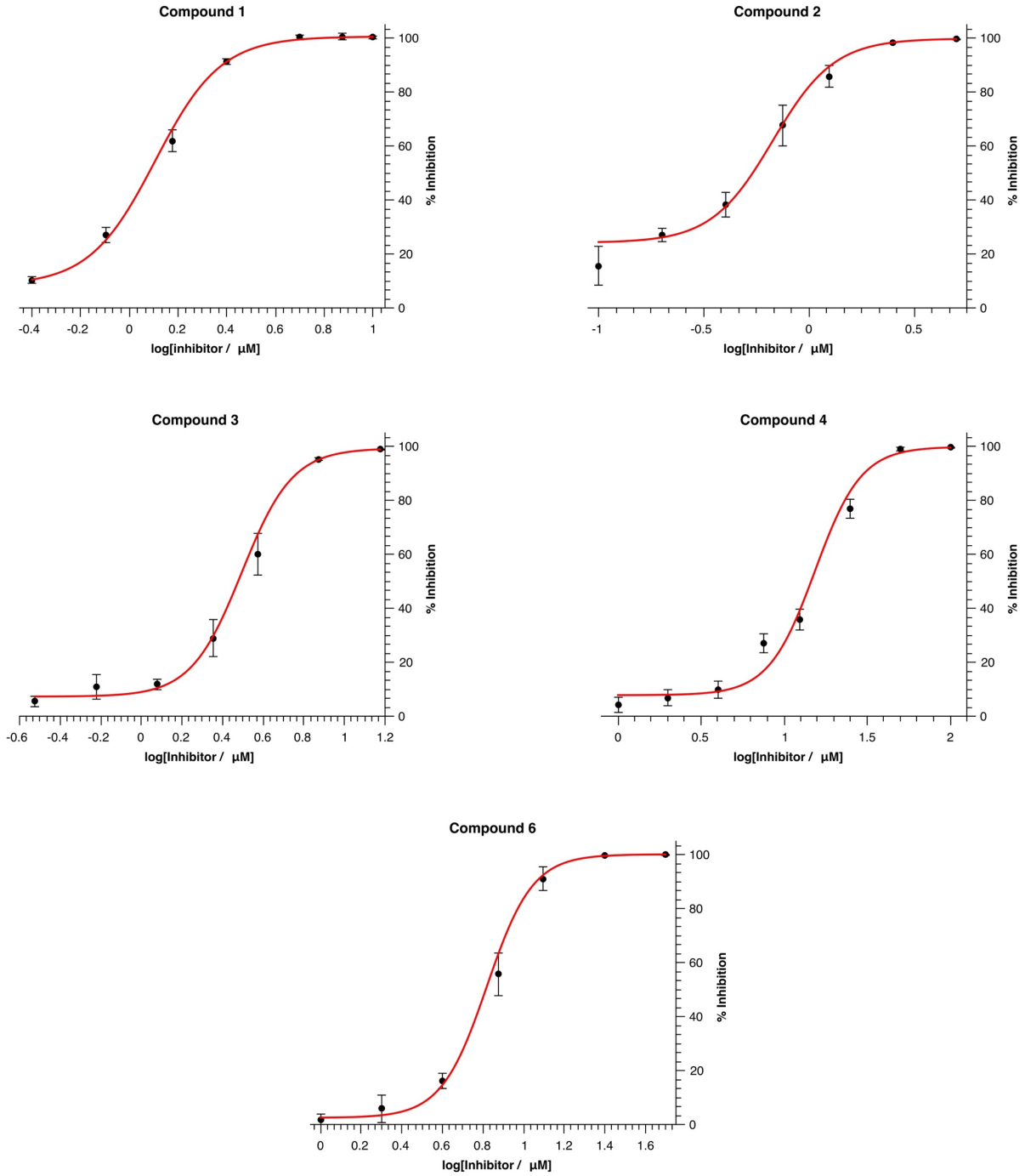


Figure S5. Dose response curves of Zika virus NS2B-NS3 protease inhibitors. Measurements were performed in triplicate at concentrations of 15 µM substrate and 1 nM linked Zika virus NS2B-NS3 protease (gZiPro). Dose-response curves were fitted using a sigmoid fitting function. Similar dose-response curves were recorded for at least three additional substrate concentrations to derive K_i values.

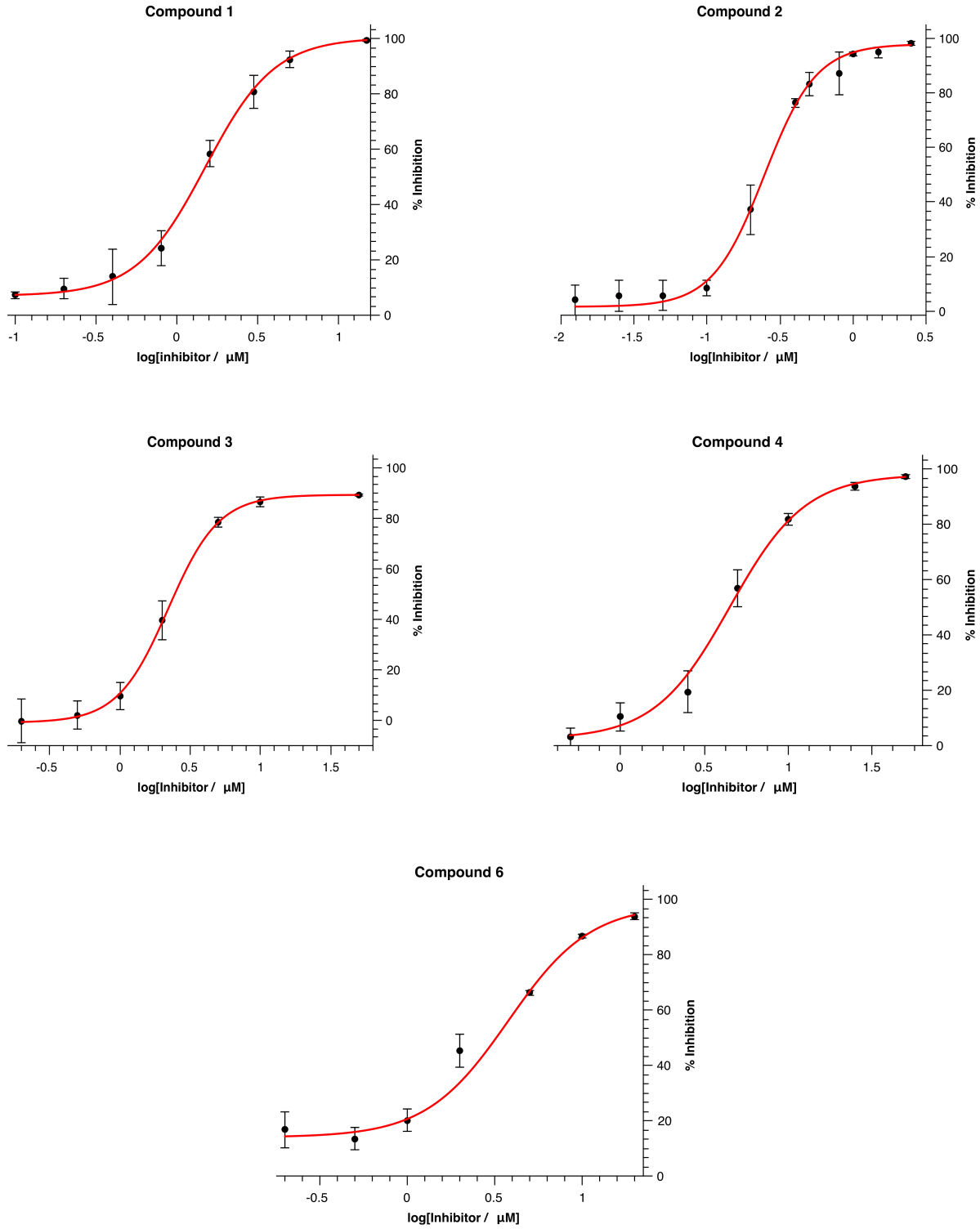


Figure S6. Dose response curves of Zika virus NS2B-NS3 protease inhibitors. Measurements were performed in triplicate at concentrations of 15 µM substrate and 0.5 nM unlinked Zika virus NS2B-NS3 protease (bZiPro). Dose-response curves were fitted using a sigmoid fitting function.

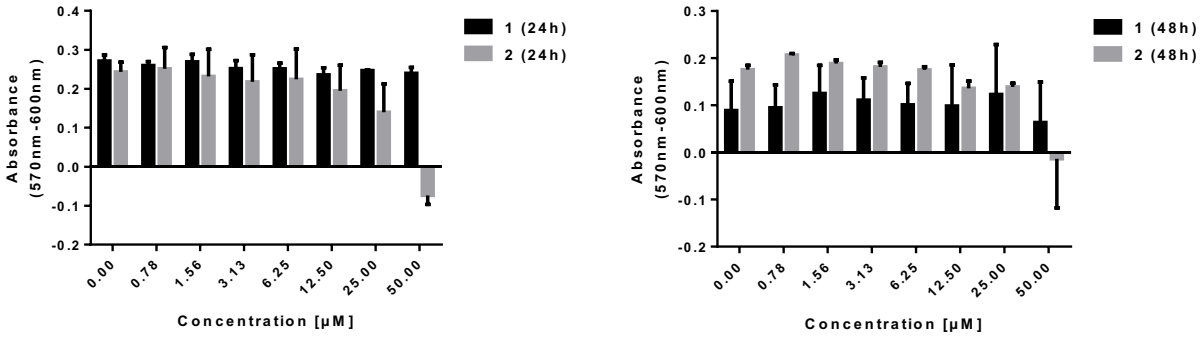


Figure S7. Absorbance data of Huh-7 cell viability assay for compounds 1 and 2. Cell viability was determined using CellTiter-Blue® reagent (Promega). The difference in absorbance at 570 nm and 600 nm is plotted against the compound concentration used for Huh-7 treatment. Measurements were performed in triplicate after 24 h (left) and 48 h (right) incubation.

Table S1. Kinetic parameters for Zika virus NS2B-NS3 protease binding of compounds identified.

Compound	k_{on} ($M^{-1} \cdot s^{-1}$)	k_{off} (s^{-1})	K_D (nM)
1	1.47 x 10^5	0.003	20
2	2.22 x 10^5	0.002	8.7
3	3.54 x 10^5	0.003	7.5
4	1.83 x 10^5	0.002	8.9
5	3.22 x 10^5	0.002	5.0
6*	3.04 x 10^4 0.004	0.042 0.001	168

* Data of compound **6** were fitted using a two-state binding model. Therefore, two values are reported for k_{on} and k_{off}.

Table S2. DNA oligonucleotides used in this work.

Name	Sequence (5' to 3')[a]
NNS-GT3-N8.R72	CTACGTTCCCGTTCCCGTTCCGCASNNSNNSNNSNNSNNSNNSNNCATATGTATATCTCCTTCTTAAAG
NNS-GT3-N9.R75	CTACGTTCCCGTTCCCGTTCCGCASNNSNNSNNSNNSNNSNNSNNSNNCATATGTATATCTCCTTCTTAAAG
NNS-GT3-N10.R78	CTACGTTCCCGTTCCCGTTCCGCASNNSNNSNNSNNSNNSNNSNNSNNSNNCATATGTATATCTCCTTCTTAAAG
T7g10M.F46	TAATACGACTCACTATAGGGTTAACTTTAAGAAGGAGATATACATA
CGT3an13.R23	TTTCCGCCCCCCGTCCTACGTTC
CGT3an13.R39	TTTCCGCCCCCCGTCCTACGTTCCCGTTCCCGTTCCGCA
EnGluCAU.R44	TCGAACCCCTGTTACCGCCTTATGAGGGCGGTGTCCTGGGCCTC
EnGluGUC.R44	TCGAACCCCTGTTACCGCCTTGACAGGGCGGTGTCCTGGGCCTC
EnGluCUC.R44	TCGAACCCCTGTTACCGCCTTGAGAGGGCGGTGTCCTGGGCCTC
EnGluGCG.R44	TCGAACCCCTGTTACCGCCTTCGCAGGGCGGTGTCCTGGGCCTC
EnGluCCG.R44	TCGAACCCCTGTTACCGCCTTCGGAGGGCGGTGTCCTGGGCCTC
EnGluCCU.R44	TCGAACCCCTGTTACCGCCTTAGGAGGGCGGTGTCCTGGGCCTC
Glu-5'.F49	GTAATACGACTCACTATAGTCCCCTTCGTCTAGAGGCCCAGGACACCGC
Glu-3'.R37	TGGCGTCCCCTAGGGGATTCGAACCCCTGTTACCGCC
Glu3'.R20(OMe)	TrG*GCGTCCCCTAGGGGATTC
Ini-3'.R20(OMe)	TrG*GTTGCGGGGGCCGGATTT
Ini-3'.R38	TGGTTGCGGGGGCCGGATTTGAACCGACGATCTTCGGG
Ini1-1G-5'.F49	GTAATACGACTCACTATAGGCGGGGTGGAGCAGCCTGGTAGCTCGTCGG
Ini cat.R44	GAACCGACGATCTTCGGGTTATGAGCCCGACGAGCTACCAGGCT
T7ex5.F22	GGCGTAATACGACTCACTATAG
eFx.R45	ACCTAACGCTAATCCCCTTTCGGGGCCGCGGAAATCTTTCGATCC
eFx.R18	ACCTAACGCTAATCCCCT
dFx.R46	ACCTAACGCCATGTACCCTTTCGGGGATGCGGAAATCTTTCGATCC
dFx.R19	ACCTAACGCCATGTACCCT

[a] rG* indicates 2'-*O*-methyl guanosine

Table S3. Non-linear least-squares fitting parameters for gZiPro (Graphpad Prism 7.0).

Compound	Non-competitive model[a]		Competitive model[b]		Mixed model[c]		
	K_i (95% CI)[d]	R^2	K_i (95% CI)[d]	R^2	K_i (95% CI)[d]	α[e] (95% CI)[d]	R^2
1	0.81 (0.68 – 0.96)	0.943	~ $1.89 \cdot 10^{-8}$ (very wide)	0.765	~ 1.47 (very wide)	~ 0.54 (very wide)	0.881
2	0.44 (0.39 – 0.50)	0.963	~ $1.26 \cdot 10^{-9}$ (very wide)	0.848	~ 0.254 (very wide)	~ 1.74 (very wide)	0.940
3	2.00 (1.65 – 2.43)	0.921	~ $2.20 \cdot 10^{-8}$ (very wide)	0.811	~ 137 (very wide)	~ 0.014 (very wide)	0.860
4	20.4 (16.9 – 24.9)	0.913	~ $5.27 \cdot 10^{-7}$ (very wide)	0.799	~ 8.75 (very wide)	~ 2.35 (very wide)	0.831
6	3.51 (2.92 – 4.24)	0.934	~ $1.07 \cdot 10^{-7}$ (very wide)	0.745	~ 7.79 (very wide)	~ 0.44 (very wide)	0.869

[a] Fitted using $v_{max}^{app} = \frac{v_{max}}{1+\frac{[I]}{K_i}}$

[b] Fitted using $K_m^{app} = K_m\left(1+\frac{[I]}{K_i}\right)$

[c] Fitted using $v_{max}^{app} = \frac{v_{max}}{1+\frac{[I]}{\alpha K_i}}$, $K_m^{app} = K_m\left(\frac{1+\frac{[I]}{K_i}}{1+\frac{[I]}{\alpha K_i}}\right)$

[d] Values for the 95% confidence interval are shown in parenthesis.

[e] Degree to which the binding of inhibitor changes the affinity of the enzyme for substrate.

References

1. Mahawaththa, M. C.; Pearce, B. J. G.; Szabo, M.; Graham, B.; Klein, C. D.; Nitsche, C.; Otting, G. Solution conformations of a linked construct of the Zika virus NS2B-NS3 protease. *Antiviral Res.* **2017,** *142*, 141-147.
2. Nitsche, C.; Mahawaththa, M. C.; Becker, W.; Huber, T.; Otting, G. Site-selective tagging of proteins by pnictogen-mediated self-assembly. *Chem. Commun.* **2017,** *53*, 10894-10897.
3. Neylon, C.; Brown, S. E.; Kralicek, A. V.; Miles, C. S.; Love, C. A.; Dixon, N. E. Interaction of the *Escherichia coli* replication terminator protein (Tus) with DNA: A model derived from DNA-binding studies of mutant proteins by surface plasmon resonance. *Biochemistry* **2000,** *39*, 11989-11999.
4. 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.
5. Goto, Y.; Ohta, A.; Sako, Y.; Yamagishi, Y.; Murakami, H.; Suga, H. Reprogramming the translation initiation for the synthesis of physiologically stable cyclic peptides. *ACS Chem. Biol.* **2008,** *3*, 120-129.
6. Hayashi, Y.; Morimoto, J.; Suga, H. *In vitro* selection of anti-Akt2 thioether-macrocyclic peptides leading to isoform-selective inhibitors. *ACS Chem. Biol.* **2012,** *7*, 607-613.
7. Jongkees, S. A. K.; Caner, S.; Tysoe, C.; Brayer, G. D.; Withers, S. G.; Suga, H. Rapid discovery of potent and selective glycosidase-inhibiting *de novo* peptides. *Cell Chem. Biol.* **2017,** *24*, 381-390.
8. Matsunaga, Y.; Bashiruddin, N. K.; Kitago, Y.; Takagi, J.; Suga, H. Allosteric inhibition of a semaphorin 4D receptor plexin B1 by a high-affinity macrocyclic peptide. *Cell Chem. Biol.* **2016,** *23*, 1341-1350.
9. Passioura, T.; Watashi, K.; Fukano, K.; Shimura, S.; Saso, W.; Morishita, R.; Ogasawara, Y.; Tanaka, Y.; Mizokami, M.; Sureau, C.; Suga, H.; Wakita, T. *De novo* macrocyclic peptide inhibitors of hepatitis B virus cellular entry. *Cell Chem. Biol.* **2018,** *25*, 906-915.
10. Song, X.; Lu, L. Y.; Passioura, T.; Suga, H. Macrocyclic peptide inhibitors for the protein-protein interaction of Zaire Ebola virus protein 24 and karyopherin alpha 5. *Org. Biomol. Chem.* **2017,** *15*, 5155-5160.
11. Yamagishi, Y.; Shoji, I.; Miyagawa, S.; Kawakami, T.; Katoh, T.; Goto, Y.; Suga, H. Natural product-like macrocyclic N-methyl-peptide inhibitors against a ubiquitin ligase

uncovered from a ribosome-expressed *de novo* library. *Chem. Biol.* **2011,** *18*, 1562-1570.

12. Goto, Y.; Katoh, T.; Suga, H. Flexizymes for genetic code reprogramming. *Nat. Protoc.* **2011,** *6*, 779-790.
13. 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.
14. Becker, W.; Adams, L. A.; Graham, B.; Wagner, G. E.; Zangger, K.; Otting, G.; Nitsche, C. Trimethylsilyl tag for probing protein-ligand interactions by NMR. *J. Biomol. NMR* **2018,** *70*, 211-218.
15. Nitsche, C.; Schreier, V. N.; Behnam, M. A.; Kumar, A.; Bartenschlager, R.; Klein, C. D. Thiazolidinone-peptide hybrids as dengue virus protease inhibitors with antiviral activity in cell culture. *J. Med. Chem.* **2013,** *56*, 8389-8403.
16. 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.