Crystal structure of unlinked NS2B-NS3 protease from Zika virus

Zhang, Zhenzhen; Li, Yan; Loh, Ying Ru; Phoo, Wint Wint; Hung, Alvin W.; Kang, CongBao;
Luo, Dahai

2016

Zhang, Z., Li, Y., Loh, Y. R., Phoo, W. W., Hung, A. W., Kang, C., et al. (2016). Crystal structure of unlinked NS2B-NS3 protease from Zika virus. *Science*, 354(6319), 1597-1600.

<https://hdl.handle.net/10356/84352>

<https://doi.org/10.1126/science.aai9309>

© 2016 The Author(s). This is the author created version of a work that has been peer reviewed and accepted for publication in *Science*, published by American Association for the Advancement of Science on behalf of The Author(s). It incorporates referee's comments but changes resulting from the publishing process, such as copyediting, structural formatting, may not be reflected in this document. The published version is available at: [<http://dx.doi.org/10.1126/science.aai9309>].

Downloaded on 26 Apr 2023 19:31:55 SGT

Crystal Structure of Unlinked NS2B-NS3 Protease from Zika Virus

Zhenzhen Zhang^{1,2,†}, Yan Li^{3, †}, Ying Ru Loh³, Wint Wint Phoo^{1,2,4}, Alvin W. Hung³, CongBao Kang^{3*}, Dahai Luo^{1,2*}

Affiliations:

¹Lee Kong Chian School of Medicine, Nanyang Technological University, EMB 03-07, 59 Nanyang Drive, Singapore 636921

²NTU Institute of Structural Biology, Nanyang Technological University, EMB 06-01, 59 Nanyang Drive, Singapore 636921.

³Experimental Therapeutics Centre, Agency for Science, Technology and Research (A*STAR), 31 Biopolis way, Nanos, #03-01, Singapore 138669.

⁴School of Biological Sciences, Nanyang Technological University, 60 Nanyang Drive, Singapore 636921.

*Correspondence to: CongBao Kang, cbkang@etc.a-star.edu.sg ; Dahai Luo, luodahai@ntu.edu.sg

† These two authors contributed equally to the work;

Abstract:

Zika virus (ZIKV) has rapidly emerged as a global public health concern. Viral NS2B-NS3 protease processes viral polyprotein and is essential for the virus replication, making it an attractive antiviral drug target. We report crystal structures of the unlinked NS2B-NS3 protease from ZIKV as free enzyme and bound to a peptide reversely oriented at the active site at 1.58 Å resolution. The unlinked NS2B-NS3 protease adopts a closed conformation in which NS2B engages NS3 to form an empty substrate binding site. A second protease in the same crystal binds to the residues K₁₄K₁₅G₁₆E₁₇ from the neighboring NS3 in reverse orientation resisting proteolysis. These features of ZIKV NS2B-NS3 protease may accelerate structure-based antiviral drug discovery against ZIKV and related pathogenic flaviviruses.

One Sentence Summary:

Separately expressed NS2B and NS3 of Zika virus forms an active enzyme that adopts a closed conformation in the crystal structure and in solution.

1 Main Text

2 Zika virus (ZIKV) has spread across the world rapidly and is becoming a serious public
3 health concern owing to its link to severe neurological diseases such as fetal microcephaly and
4 Guillain-Barré syndrome in adults (1, 2). Specific antiviral therapeutics against ZIKV are
5 urgently needed to fight this pandemic. ZIKV belongs to the flaviviridae family, flavivirus
6 genus, which contains important human pathogens including dengue virus (DENV), West Nile
7 virus (WNV), yellow fever virus (YFV), Japanese encephalitis virus (JEV), and tick-borne
8 encephalitis virus (TBEV) (3-5). The genome of these viruses encodes a polyprotein that is
9 processed into three structural proteins (capsid, membrane, and envelope proteins) and seven
10 nonstructural (NS) proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5) by both host
11 proteases and the viral NS2B-NS3 protease. As such, the NS2B-NS3 protease is an attractive
12 target for antiviral drug development (6, 7). NS3 contains a trypsin-like fold and carries the
13 conserved catalytic triad (H₅₁, D₇₅ and S₁₃₅). The small trans-membrane protein NS2B anchors
14 NS3 to the endoplasmic reticulum membrane, and together they form an active enzyme for
15 substrate recognition and catalysis (7-10). The minimal cofactor region of the NS2B comprises
16 the hydrophilic residues 49-97 (11). The N-terminal 18 residues (49-67) of NS2B cofactor
17 support the proper fold of NS3 protease by forming a beta strand inserted into the protease
18 domain (12, 13). The C-terminal part of NS2B cofactor (residues 68-96) forms a β -hairpin to
19 create the S2 and S3 pockets in the substrate-binding site of NS3pro (13-15).

20 Crystal structures of the proteases from DENV (13, 14, 16), WNV(13, 15, 17, 18),
21 Murray Valley encephalitis virus (MVEV)(19) and ZIKV(20) have been determined using the
22 same construct of NS2B-NS3 connected with an artificial glycine-rich linker. In the absence of
23 an inhibitor, the C-terminal part of NS2B cofactor is flexible and often disordered ("open

conformation")(13, 16, 18, 19). Inhibitor-bound structures adopt a compact complex where the NS2B fragment wraps around the NS3 protease and makes direct contacts with the inhibitors ("closed conformation") (13-15, 17-20). While solution NMR studies on DENV and WNV proteases suggested that the free enzyme is able to form the closed conformation in solution (21-23), it is still unclear whether the closed conformation is stabilized by the binding of a ligand. For ZIKV, we have discovered that the artificial linker introduces steric hindrance and alters the substrate (inhibitor) binding behavior, suggesting that the unlinked binary ZIKV NS2B:NS3 protease (bZiPro) is preferable for studying the enzyme behavior and for inhibitor development(24). In this regard, it is important to characterize the dynamic behavior of ZIKV NS2B-NS3 protease in great detail.

Here we report the crystal structure of bZiPro at a resolution of 1.58 Å (**Fig. 1 and Table S1**). The refined model consists of four bZiPro molecules in one unit cell labeled according to the peptide chain IDs *A/B*, *C/D*, *E/F*, *G/H* (**Fig. S1**). There are three molecules of free enzyme and one bound to K₁₄K₁₅G₁₆E₁₇ tetrapeptide from the neighboring NS3 N-terminus in an unusual reverse orientation (**Fig. 1 and Fig. S1**). NS2B cofactor is in the closed conformation as both N- and C-terminal regions are folded into a β-sheet conformation and in a complex with the NS3 protease domain in all four bZiPro molecules. The structures of bZiPro free enzyme are virtually identical to the peptide-bound bZiPro (*AB*) with RMSD values of 0.25 to 0.34 Å after superimposition (**Fig. 1D, Fig. S1B and Table S2**). In addition, the temperature factor of individual NS2B chain is also comparable to that of the partner NS3, indicating that the NS2B and NS3 form a stable complex in the crystal (**Fig. S1C**). Furthermore, bZiPro is also very similar to eZiPro (ZIKV NS2B-NS3 protease after self-cleavage; PDB code: 5GJ4) (RMSD value of 0.45 Å for 162 Cα atoms) and gZiPro (the single chain ZIKV NS2B₄₉₋₉₆-G₄SG₄-NS3Pro

1 for NS2B-NS3 protease with a glycine-rich linker; PDB code: 5LC0) (RMSD value of 0.52 Å for
2 167 Cα atoms) (**Fig. 1, Fig. S1A and Table S2**)(20, 24). The unlinked ZIKV NS2B-NS3
3 protease appears to contain a preformed stable substrate binding pocket which does not undergo
4 further significant conformational changes upon substrate or inhibitor binding.

5 Surprisingly, one bZiPro molecule (*AB*) binds to the N terminal K₁₄K₁₅G₁₆E₁₇
6 tetrapeptide sequence extended from the neighbor NS3 protease (chain H') (**Fig. 1A and Fig.**
7 **S1A**). We identified the residues E₁₂-T₁₈ from the neighboring NS3 N-terminal region in the
8 electron density map and built it into the structure unambiguously (**Fig. 1F**). The same residues
9 are disordered in the remaining three bZiPro, eZiPro and gZiPro structures, suggesting that the
10 AB:H' protease complex might be a crystallographic artifact(20, 24). Nonetheless, the structure
11 does indeed capture the protease in complex with a reverse peptide. The tetra-peptide K₁₄K₁₅G
12 ₁₆E₁₇ folds into a small hairpin loop to occupy the active site. Specifically, K₁₄ ε-amino group
13 occupies the S1 pocket and forms hydrogen bonds with D₁₂₉ and Y₁₃₀. Residue K₁₅ contributes
14 the most to the overall binding: its side chain ε-amino group forms hydrogen bonds with S₈₁ and
15 D₈₃ in the S2 pocket and its main chain forms hydrogen bonds with G₁₅₁ and G₁₅₃ of NS3
16 protease. G₁₆ does not form any interactions with the protease and leaves the S3 pocket
17 completely empty, similar to NS2B G₁₂₉ in the eZiPro structure. E₁₇ side chain stacks with the
18 aromatic ring of Y₁₆₁ of the protease and partially occupies the S1 pocket with K₁₄. Notably, the
19 hairpin is partially stabilized by the intra-molecular hydrogen bonds: one between the backbone
20 carbonyl oxygen atom of K₁₄ and amide nitrogen atom of E₁₇; two between the side chain
21 carboxylic group of E₁₇ and the ε-amino group of K₁₄ and backbone amide of T₁₈ (**Fig. 1D**). The
22 specific conformation of the reverse peptide does not allow the catalytic residue S₁₃₅ to establish
23 the hydrogen bonding relay and form the reactive oxyanion species. It also prevents the carbonyl

group of the peptide to position close to S₁₃₅. This reverse peptide could not be cleaved by the protease and the complex structure may guide structure-based inhibitor design (25).

To reveal the structural dynamics of bZiPro in solution, we carried out NMR studies. The acquired ¹H-¹⁵N-HSQC spectrum of bZiPro exhibited dispersed cross-peaks, indicating that bZiPro is a well folded structure with similar secondary structures to those of bZiPro crystal structure (**Fig. 2A** and **Fig. S5**). Due to signal broadening in ¹H-¹⁵N-HSQC spectrum of bZiPro, there are missing peaks for residues 47-49, 114-119, 123-133, 153-158, and 161-166 of NS3 (**Fig. 2B, 2C** and **Fig. S4B**). To determine whether the protease exists in the closed conformation in solution with an open active site for inhibitor binding, we titrated bZiPro with acetyl-lysine-arginine (AcKR) (**Fig. 2**), a dipeptide known to weakly inhibit WNV protease activity (IC₅₀ ≥ 100 μM) (26). The ¹H-¹⁵N-HSQC spectra of bZiPro in complex with AcKR are better resolved than bZiPro alone and cross-peaks of many residues from both NS2B and NS3 re-appeared (**Fig. 2** and **Fig. S4**). ¹H-¹⁵N-HSQC profile of bZiPro:AcKR is very similar to that of eZiPro, implying that the cleavage product peptides can bind to the protease active site similarly, in *cis* or *trans* (**Fig. S6**). As seen from the ¹H-¹⁵N-HSQC profile, residues from the catalytic triad and their surroundings undergo local environmental changes upon peptide binding (**Fig. S7**). H₅₁ first exhibited line broadening in the presence of 0.4-1.6 mM AcKR peptide. Its peak reappeared at a different position when excess amount of inhibitor (3.2-6.4 mM) was used to saturate the active site. Cross peaks of NS3 residues, such as 114-119, 125-130, 133-134, 153-158, 161 and 163-166 and the C terminal region of NS2B cofactor also appeared as a β-hairpin (**Fig 2** and **Fig. S5**). We also observed a second population of narrowly dispersed cross-peaks of unstructured NS2B present in both bZiPro and bZiPro-AcKR complex spectra. These peaks were identical to that of the free NS2B cofactor and were unresponsive to AcKR titration (**Fig. S4** and

Fig. S5). There seems to be free NS2B cofactor after dissociation from NS3 present in the bZiPro protein samples.

The overall flexibility changes of bZiPro upon binding to AcKR were further probed by ^{15}N -R₁, R₂ and heteronuclear NOE (hetNOE) experiments (**Fig. 2E** and **Fig. S8**). Consistent with previous structural studies, the N-terminal part of NS2B cofactor exhibited very similar dynamic patterns to those of NS3 independent of AcKR binding as this segment of NS2B forms a very stable complex with NS3. The C-terminal β -hairpin region also forms stable structure in solution, evidenced by the high hetNOE values. Taken together, bZiPro displayed a very dynamic local environment at both the active site and the interface between the C-terminal part of NS2B cofactor and NS3, which is stabilized by addition of a peptide inhibitor occupying both S1 and S2 pockets (**Fig. 2**). The free bZiPro adopts a closed conformation in solution and the line broadening observed for the residues may be due to presence of minor local conformational exchanges of the C-terminus of NS2B and residues close to the active site of NS3. Similarly, it has been reported that minor populations or local conformational exchanges can lead to disappearance of cross peaks for WNV protease (22).

To prove that the bZiPro construct is suitable for antiviral drug design, we determined the crystal structure of bZiPro in complex with a fragment EN300 (chemical name: 1H-benzo[d]imidazol-1-yl)methanol)(**Fig. 3**). The compound was identified through a fragment screen and was found to be able to stabilize bZiPro in a thermal shift assay (**Fig. S9**). Changes of the ^1H - ^{15}N -HSQC spectra upon the compound binding not only confirmed the direct binding of the compound to bZiPro but also showed that the compound did not cause any large scale conformational changes to the protein (**Fig. 3A**). Indeed, the compound EN300 is sandwiched between Y₁₆₁ aromatic ring and A₁₃₂ via stacking interactions and forms a hydrogen bond with

Y₁₅₀ (**Fig. 3**). The compound only occupies the S1 pocket partially and forms no direct contact with NS2B cofactor. Therefore the closed conformation of bZiPro is captured again in this structure, which is unlikely due to the binding of the small compound. The protein-ligand complex structure serves as a starting point to guide further chemical modifications for optimizations in binding potency and inhibition of protease activity inhibition in a targeted manner.

In conclusion, we show that the crystal structure of free unlinked ZIKV NS2B-NS3 protease exists in the closed conformation. Solution NMR studies confirmed that free protease is predominantly in a closed conformation while local conformational dynamics occurs at the NS2B-NS3 binding interface. Substrate peptide binding does not further induce significant conformational changes. This finding has relevance for antiviral drug discovery targeting ZIKV protease. The new crystal form at high resolution presented in our work will be of practical use because these crystals can be readily soaked or co-crystallized with inhibitors. Chemical library screens including fragment-based screening will have a higher likelihood of obtaining potent lead compounds and successful structure-based lead optimization. We also report the co-crystal structure of a flaviviral protease binding to a peptide in a C-to-N orientation. The reverse direction of the peptide bond is unable to form tetrahedral intermediate at the protease active site and is therefore not cleavable. Although this might be a crystal artifact, the good fit of the peptide within the protease pocket provides an attractive starting point for developing novel peptidomimetic inhibitors, for instance cyclic-peptides.

References:

1. V. M. Cao-Lormeau, A. Blake, S. Mons, S. Lastere, C. Roche, J. Vanhomwegen, T. Dub, L. Baudouin, A. Teissier, P. Larre, A. L. Vial, C. Decam, V. Choumet, S. K. Halstead, H. J. Willison, L. Musset, J. C. Manuguerra, P. Despres, E. Fournier, H. P. Mallet, D. Musso, A. Fontanet, J. Neil, F. Ghawche, Guillain-Barre Syndrome outbreak associated with Zika virus infection in French Polynesia: a case-control study. *Lancet* **387**, 1531-1539 (2016).
2. L. Schuler-Faccini, E. M. Ribeiro, I. M. Feitosa, D. D. Horovitz, D. P. Cavalcanti, A. Pessoa, M. J. Doriqui, J. I. Neri, J. M. Neto, H. Y. Wanderley, M. Cernach, A. S. El-Husny, M. V. Pone, C. L. Sero, M. T. Sanseverino, F. Brazilian Medical Genetics Society-Zika Embryopathy Task, Possible Association Between Zika Virus Infection and Microcephaly - Brazil, 2015. *MMWR. Morbidity and mortality weekly report* **65**, 59-62 (2016).
3. T. C. Pierson, M. S. Diamond, in *Fields Virology. 6th Edition*. (Philadelphia: Lippincott-Raven Publishers., 2013), chap. 26, pp. 747-794.
4. B. D. Lindenbach, M. C.L., H. J. Thiel, C. M. Rice, in *Fields Virology. 6th Edition*. (Philadelphia: Lippincott-Raven Publishers., 2013), chap. 25, pp. 712-746.
5. S. C. Weaver, F. Costa, M. A. Garcia-Blanco, A. I. Ko, G. S. Ribeiro, G. Saade, P. Y. Shi, N. Vasilakis, Zika virus: History, emergence, biology, and prospects for control. *Antiviral Res* **130**, 69-80 (2016).
6. S. Ekins, D. Mitchen, M. Coffee, T. P. Stratton, J. S. Freundlich, L. Freitas-Junior, E. Muratov, J. Siqueira-Neto, A. J. Williams, C. Andrade, Open drug discovery for the Zika virus. *F1000Research* **5**, 150 (2016).
7. C. Nitsche, S. Holloway, T. Schirmeister, C. D. Klein, Biochemistry and medicinal chemistry of the dengue virus protease. *Chemical reviews* **114**, 11348-11381 (2014).
8. D. Luo, S. G. Vasudevan, J. Lescar, The flavivirus NS2B-NS3 protease-helicase as a target for antiviral drug development. *Antiviral Res* **118**, 148-158 (2015).
9. B. Falgout, M. Pethel, Y. M. Zhang, C. J. Lai, Both nonstructural proteins NS2B and NS3 are required for the proteolytic processing of dengue virus nonstructural proteins. *J Virol* **65**, 2467-2475 (1991).
10. T. J. Chambers, R. C. Weir, A. Grakoui, D. W. McCourt, J. F. Bazan, R. J. Fletterick, C. M. Rice, Evidence that the N-terminal domain of nonstructural protein NS3 from yellow fever virus is a serine protease responsible for site-specific cleavages in the viral polyprotein. *Proc Natl Acad Sci U S A* **87**, 8898-8902 (1990).
11. D. Leung, K. Schroder, H. White, N. X. Fang, M. J. Stoermer, G. Abbenante, J. L. Martin, P. R. Young, D. P. Fairlie, Activity of recombinant dengue 2 virus NS3 protease in the presence of a truncated NS2B co-factor, small peptide substrates, and inhibitors. *J Biol Chem* **276**, 45762-45771 (2001).
12. D. Luo, T. Xu, C. Hunke, G. Gruber, S. G. Vasudevan, J. Lescar, Crystal structure of the NS3 protease-helicase from dengue virus. *J Virol* **82**, 173-183 (2008).
13. P. Erbel, N. Schiering, A. D'Arcy, M. Renatus, M. Kroemer, S. P. Lim, Z. Yin, T. H. Keller, S. G. Vasudevan, U. Hommel, Structural basis for the activation of flaviviral NS3 proteases from dengue and West Nile virus. *Nat Struct Mol Biol* **13**, 372-373 (2006).
14. C. G. Noble, C. C. Seh, A. T. Chao, P. Y. Shi, Ligand-bound structures of the dengue virus protease reveal the active conformation. *J Virol* **86**, 438-446 (2012).

- 1 15. G. Robin, K. Chappell, M. J. Stoermer, S. H. Hu, P. R. Young, D. P. Fairlie, J. L. Martin,
2 Structure of West Nile virus NS3 protease: ligand stabilization of the catalytic
3 conformation. *J Mol Biol* **385**, 1568-1577 (2009).
- 4 16. S. Chandramouli, J. S. Joseph, S. Daudenarde, J. Gatchalian, C. Cornillez-Ty, P. Kuhn,
5 Serotype-specific structural differences in the protease-cofactor complexes of the dengue
6 virus family. *J Virol* **84**, 3059-3067 (2010).
- 7 17. M. Z. Hammamy, C. Haase, M. Hammami, R. Hilgenfeld, T. Steinmetzer, Development
8 and characterization of new peptidomimetic inhibitors of the West Nile virus NS2B-NS3
9 protease. *ChemMedChem* **8**, 231-241 (2013).
- 10 18. A. E. Aleshin, S. A. Shiryaev, A. Y. Strongin, R. C. Liddington, Structural evidence for
11 regulation and specificity of flaviviral proteases and evolution of the Flaviviridae fold.
12 *Protein science : a publication of the Protein Society* **16**, 795-806 (2007).
- 13 19. R. Assenberg, E. Mastrangelo, T. S. Walter, A. Verma, M. Milani, R. J. Owens, D. I.
14 Stuart, J. M. Grimes, E. J. Mancini, Crystal structure of a novel conformational state of the
15 flavivirus NS3 protein: implications for polyprotein processing and viral replication. *J*
16 *Virol* **83**, 12895-12906 (2009).
- 17 20. J. Lei, G. Hansen, C. Nitsche, C. D. Klein, L. Zhang, R. Hilgenfeld, Crystal structure of
18 Zika virus NS2B-NS3 protease in complex with a boronate inhibitor. *Science* **353**, 503-505
19 (2016).
- 20 21. L. de la Cruz, W. N. Chen, B. Graham, G. Otting, Binding mode of the activity-modulating
21 C-terminal segment of NS2B to NS3 in the dengue virus NS2B-NS3 protease. *FEBS J* **281**,
22 1517-1533 (2014).
- 23 22. X. C. Su, K. Ozawa, R. Qi, S. G. Vasudevan, S. P. Lim, G. Otting, NMR analysis of the
24 dynamic exchange of the NS2B cofactor between open and closed conformations of the
25 West Nile virus NS2B-NS3 protease. *PLoS neglected tropical diseases* **3**, e561 (2009).
- 26 23. Y. M. Kim, S. Gayen, C. Kang, J. Joy, Q. Huang, A. S. Chen, J. L. Wee, M. J. Ang, H. A.
27 Lim, A. W. Hung, R. Li, C. G. Noble, T. Lee le, A. Yip, Q. Y. Wang, C. S. Chia, J. Hill, P.
28 Y. Shi, T. H. Keller, NMR analysis of a novel enzymatically active unlinked dengue
29 NS2B-NS3 protease complex. *J Biol Chem* **288**, 12891-12900 (2013).
- 30 24. W. W. Phoo, Y. Li, Z. Zhang, M. Y. Lee, Y. R. Loh, Y. B. Tan, E. Y. Ng, J. Lescar, C.
31 Kang, D. Luo, Structure of the NS2B-NS3 protease from Zika virus after self-cleavage.
32 *Nature communications* **7**, 13410 (2016).
- 33 25. C. Nitsche, M. A. Behnam, C. Steuer, C. D. Klein, Retro peptide-hybrids as selective
34 inhibitors of the Dengue virus NS2B-NS3 protease. *Antiviral Res* **94**, 72-79 (2012).
- 35 26. C. Kang, S. Gayen, W. Wang, R. Severin, A. S. Chen, H. A. Lim, C. S. Chia, A. Schuller,
36 D. N. Doan, A. Poulsen, J. Hill, S. G. Vasudevan, T. H. Keller, Exploring the binding of
37 peptidic West Nile virus NS2B-NS3 protease inhibitors by NMR. *Antiviral Res* **97**, 137-
38 144 (2013).
- 39 27. K. Pervushin, A. Ono, C. Fernandez, T. Szyperski, M. Kainosho, K. Wuthrich, NMR scalar
40 couplings across Watson-Crick base pair hydrogen bonds in DNA observed by transverse
41 relaxation-optimized spectroscopy. *Proc Natl Acad Sci U S A* **95**, 14147-14151 (1998).
- 42 28. M. Salzmann, K. Pervushin, G. Wider, H. Senn, K. Wuthrich, TROSY in triple-resonance
43 experiments: new perspectives for sequential NMR assignment of large proteins. *Proc Natl*
44 *Acad Sci U S A* **95**, 13585-13590 (1998).

29. F. Delaglio, S. Grzesiek, G. W. Vuister, G. Zhu, J. Pfeifer, A. Bax, NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* **6**, 277-293 (1995).
30. B. A. Johnson, Using NMRView to visualize and analyze the NMR spectra of macromolecules. *Methods Mol Biol* **278**, 313-352 (2004).
31. Y. Shen, F. Delaglio, G. Cornilescu, A. Bax, TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J Biomol NMR* **44**, 213-223 (2009).
32. L. E. Kay, D. A. Torchia, A. Bax, Backbone dynamics of proteins as studied by ¹⁵N inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. *Biochemistry* **28**, 8972-8979 (1989).
33. S. Gayen, Q. Li, A. S. Chen, T. H. Nguyen, Q. Huang, J. Hill, C. Kang, An NMR study of the N-terminal domain of wild-type hERG and a T65P trafficking deficient hERG mutant. *Proteins* **79**, 2557-2565 (2011).
34. T. G. Battye, L. Kontogiannis, O. Johnson, H. R. Powell, A. G. Leslie, iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta crystallographica. Section D, Biological crystallography* **67**, 271-281 (2011).
35. P. R. Evans, An introduction to stereochemical restraints. *Acta crystallographica. Section D, Biological crystallography* **63**, 58-61 (2007).
36. P. R. Evans, An introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta crystallographica. Section D, Biological crystallography* **67**, 282-292 (2011).
37. E. Potterton, P. Briggs, M. Turkenburg, E. Dodson, A graphical user interface to the CCP4 program suite. *Acta crystallographica. Section D, Biological crystallography* **59**, 1131-1137 (2003).
38. A. Vagin, A. Teplyakov, MOLREP: an automated program for molecular replacement. *Journal of applied crystallography* **30**, 1022-1025 (1997).
39. P. D. Adams, P. V. Afonine, G. Bunkoczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, P. H. Zwart, PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-221 (2010).
40. P. Emsley, K. Cowtan, Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-2132 (2004).
41. L. Schrödinger, *The PyMOL Molecular Graphics System, Version 1.5.0.4* (2010).
42. Y. Shen, A. Bax, Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J Biomol NMR* **56**, 227-241 (2013).
43. M. P. Williamson, Using chemical shift perturbation to characterise ligand binding. *Progress in Nuclear Magnetic Resonance Spectroscopy* **73**, 1-16 (2013).

Acknowledgements. We thank W. Bin (Nanyang Technological University) and scientists from Australian Light Source MX beam-line for their help with diffraction data collection. We thank S. Liu for technical support for fragment screening. We appreciate A. Matter and T.H. Keller from for critical reading of this manuscript. The data presented in this manuscript and tabulated in the main paper and in the supplementary materials. The corresponding coordinates and structure factors are available from the PDB under accession codes 5GPI for bZiPro and 5H4I for its complex with the compound EN300. Assignments of protease in the absence and presence of AcKR have been deposited in the Biological Magnetic Resonance Data Bank (BMRB) with accession numbers 26928, 26927, respectively. This work was supported by (1) a start-up grant from Lee Kong Chian School of Medicine, Nanyang Technological University, (2) National Medical Research Council grant CBRG15May045, to DL lab, (3) A*STAR JCO grant (1431AFG102/1331A028) to CK. The authors declare no competing financial interests.

Figure Legends

Figure 1. Crystal structure of bZiPro. (A) The free-enzyme form and the peptide-bound ZIKV protease structures determined from the bZiPro crystal. In the free-enzyme form, NS2B is colored in green and NS3 in cyan. In the peptide bound form, NS2B is colored in magenta and NS3 in yellow. The N- and C-terminal residues of NS2B and NS3 are labeled. (B) Electrostatic view of the free enzyme of bZiPro with an empty pre-formed substrate binding pocket. (C) Electrostatic view of bZiPro in complex with the NS3 N-terminal peptide K₁₄K₁₅E₁₆G₁₇ in an unusual reversed orientation. Electrostatic surfaces are colored by electrostatic potential at neutral pH from -5 kT (red) to +5 kT (blue) using PyMOL. (D) Close-up views of the interactions between N-terminal residues K14-E17 from a neighbor NS3 (in cyan) and the residues from bZiPro. (E) Superimposition of the two structures of bZiPro shown as ribbons. RMSD value is 0.34 Å for 152 Cα atoms. (F) A simulated annealing omit mF_o-DF_c map of the KKGE reverse peptide within the bZiPro crystal structure is contoured at 3 σ in green mesh. 2mF_o-DF_c electron density map is contoured at 1 σ in blue.

A

B

C

D

E

F

1

2

3

Figure 2. Structural dynamics of bZiPro in solution. (A) Overlay of ^1H - ^{15}N -HSQC spectra of 0.8 mM bZiPro in the absence (black) and presence (red) of different amounts of AcKR peptide. Some residues exhibited line-broadening or chemical shift changes are labeled. (B) Residues affected by peptide binding. Enlarged view of several residues exhibited cross peak appearing and chemical shift perturbation upon AcKR binding. (C) Residues affected by AcKR binding. Affected residues shown in (B) are highlighted in spheres and labeled on the structure of protease. The structure of AcKR is modeled in the active site based on the structure of eZiPro. Residues from NS2B are underlined. Residues colored in yellow from NS3 (in magenta color for residues from NS2B) could not be unambiguously assigned due to line broadening and miss of signal connection in the free bZiPro enzyme. (D) Residues exhibited chemical shift perturbation upon addition of AcKR peptide. Residues T₅₃, K₈₄, and L₁₄₉ exhibited averaged chemical shift change more than 0.08 ppm are used for affinity estimation. Overall AcKR binds to bZiPro very weakly. (E) Protein flexibility analysis. The ^{15}N R₂ values of bZiPro in the absence and presence of AcKR peptide were plotted against residue number. The data were acquired using a 0.8 mM of bZiPro the absence (black) and presence (red) 3.2 mM of AcKR. Error bars were obtained from NMRView (30).

1

2

Figure 3. Structure of bZiPro in complex with a compound fragment. (A) Overlay of ^1H - ^{15}N -HSQC spectra of bZiPro in the absence (black) and presence (red) of EN300. (B) The simulated annealing omit map of the bZiPro crystal structure is contoured at 4σ in green mesh. The only significant peak next to the protein is evident. (C) Close-up view of the interactions between NS3 and the compound. (D) Electron density map at the compound binding site. $2mF_o - DF_c$ electron density map is contoured at 1σ in blue and the simulated annealing omit map is contoured at 3σ in green. Clear electron density is observed for the benzimidazole portion of EN300 forming a pi-stacking interaction with Y₁₆₁. Weaker electron density observed around the hydroxyl group could be indicative of partial hydrolysis of the compound.

Supplementary Materials for

Crystal Structures of Unlinked NS2B-NS3 Protease from Zika Virus

Zhenzhen Zhang^{1,2,†}, Yan Li^{3,†}, Wint Wint Phoo^{1,2,4}, Ying Ru Loh³, Alvin W. Hung³, CongBao Kang^{3*}, Dahai Luo^{1,2*}

Affiliations

¹Lee Kong Chian School of Medicine, Nanyang Technological University, EMB 03-07, 59 Nanyang Drive, Singapore 636921

²NTU Institute of Structural Biology, Nanyang Technological University, EMB 06-01, 59 Nanyang Drive, Singapore 636921.

³Experimental Therapeutics Centre, Agency for Science, Technology and Research (A*STAR), 31 Biopolis way, Nanos, #03-01, Singapore 138669.

⁴School of Biological Sciences, Nanyang Technological University, 60 Nanyang Drive, Singapore 636921.

*Correspondence to: Congbao Kang, cbkang@etc.a-star.edu.sg ; Dahai Luo, luodahai@ntu.edu.sg

† These two authors contributed equally to the work;

This PDF file includes:

Materials and Methods
Figures S1 to S9
Tables S1 to S2
Author contributions
References (23-43)

1 **Materials and Methods**

2 **Materials**

3 As described previously, the cDNA encoding NS2B cofactor (residue T45-E96) and NS3
4 protease domain (residues S1 to E177) of ZIKV (GenBank Protein Accession number
5 AEN75265.1) used in this study was synthesized (Genscript)(24). To express both NS2B
6 cofactor region and NS3 protease (bZiPro) at the same time, NS3pro (residues 1-177) and NS2B
7 cofactor region (residues 45 – 96) were cloned into the NcoI/HindIII sites, and NdeI/ XhoI sites
8 of pETDUET-1 (Novagene), respectively. To express free NS2B cofactor region, the
9 aforementioned cDNA was also cloned into pET15b. The AcKR peptides were synthesized from
10 GL Biochem Ltd (Shanghai, China) with more than 95% purity. Compound 1H-1,3-benzodiazol-
11 1-ylmethanol (EN300-17421, purchased from Enamine LTD) was dissolved in 100% DMSO to
12 prepare a 200 mM stock solution.

13 **Protein expression and purification for NMR studies**

14 The plasmid was transformed into *E. coli* BL21(DE3) competent cells and protein expression
15 was induced for overnight at 18 °C by addition of β -D-1-thiogalactopyranoside (IPTG) to 1 mM
16 final concentration, when the cell optical density at 600 nm (OD_{600}) reached 0.8-1.0. To obtain
17 ^{15}N , or ^{13}C , ^{15}N and ^2H labeled protein, cells were grown in minimal medium contain 1 g/L
18 $^{15}\text{NH}_4\text{Cl}$, or 1 g/L $^{15}\text{NH}_4\text{Cl}$, 2 g/L ^{13}C -glucose, and 100% D_2O . After induction, cells were
19 harvested by centrifugation at 11,000 $\times g$ and 4 °C for 10 min. The cells were re-suspended in a
20 buffer containing 20 mM sodium phosphate, pH 7.8, 500 mM NaCl, and 1.4 mM β -
21 mercaptoethanol. Proteins were purified using immobilized metal affinity chromatography
22 (IMAC) using Ni^{2+} -NTA resin and gel filtration chromatography as described previously using a
23 HiPrep 26/60 Sephacryl S-200 column (23). Thrombin was added to protein solution to cleave
24 the tag at the N-terminus. Cleaved proteins were further purified using gel filtration
25 chromatography in a buffer that contained 20 mM HEPES, pH7.3, 150 mM NaCl and 1 mM
26 DTT. Protein was concentrated to ~0.8 mM for data acquisition. ^{15}N -labeled NS2B cofactor
27 region alone was expressed and purified using the same protocol.

28 **Protein expression and purification for crystallography**

29 For crystallization, the following purification protocols were used: Cells were grown in LB broth
30 enriched with 50 mM potassium phosphate buffer pH 7.4 and with 2.5% glycerol. Protein
31 expression was induced overnight at 18°C by addition of IPTG to 1 mM final concentration
32 when the OD_{600} reached 0.8. Cells were harvested by centrifugation at 4000 rpm at 4°C for 25
33 min. The cells were re-suspended in lysis buffer containing 20 mM phosphate buffer saline
34 (PBS), pH 8.0, 500 mM NaCl, 2 mM β -mercaptoethanol, 10 mM imidazole and 5% glycerol.
35 Cells were lysed by passing through a LM20 microfluidizer (Microfluidics, USA) at 20,000 psi
36 and soluble fraction was obtained by centrifuging at 25,000 rpm for 1 hour. Protein was purified
37 using IMAC using Ni^{2+} -NTA beads (Qiagen, USA). After washing with lysis buffer containing
38 20-30 mM imidazole, protein was eluted with an elution buffer containing 300 mM imidazole.
39 To remove the His-tag, Thrombin was added to the eluted fractions containing the protein and
40 the mixture was dialyzed against the gel filtration buffer containing 20 mM Na-HEPES pH 7.5,
41 150 mM NaCl, 5% glycerol, 2 mM DTT for 20 hours. After dialysis, the protein was further
42 purified by size exclusion chromatography using a HiLoad 16/600 Superdex 75 pg column (GE
43 Healthcare, USA). Fractions were collected and concentrated in a final buffer containing 20 mM

1 HEPES pH 7.5, 150 mM NaCl, 5% glycerol, and 2mM DTT using spin concentrators (Vivaspin,
2 Satorius, 10 kDa MWCO).

3 **NMR measurements**

4 Backbone resonance assignment was conducted using a $^{15}\text{N}/^{13}\text{C}/^2\text{H}$ -labeled bZiPro protease in a
5 buffer that contained 20 mM Na-HEPES, 150 mM NaCl, and 1 mM DTT. Transverse relaxation-
6 optimized spectroscopy (TROSY) (27, 28)-based experiments that included ^1H - ^{15}N -HSQC, 3D-
7 HNCACB, 3D-HNCOCACB, 3D-HNCOCA, 3D-HNCA, 3D-HNCACO, 3D-HNCO and
8 NOESY-TROSY (with a mixing time of 120 ms) experiments were collected and processed by
9 TOPSPIN 2.1 (Bruker) and NMRPipe (29). All the spectra were collected on Bruker Avance II
10 700 MHz equipped with a cryoprobe at 25 °C due to the sample stability. The data were
11 processed using Topspin 2.1 and NMRPipe (29). The processed data were visualized using
12 NMRView (30) and CARA. Protein secondary structure were determined using TALOSN (31).

13
14 For the comparison of eZiPro and bZiPro in the presence of AcKR peptide, the 2D ^1H - ^{15}N -
15 HSQC spectra were acquired at 37 °C. All other titration experiments are collected at 25 °C. For
16 fragment binding experiment, 2D ^1H - ^{15}N -HSQC spectra of 0.8 mM bZiPro in the absence and
17 presence of 4.8 mM of EN300 were collected and compared.

18
19 The ^{15}N - R_1 (spin-lattice relaxation time), R_2 (spin-spin relaxation time) and steady-state
20 heteronuclear NOE (hetNOE) measurements (32) were acquired using the triple labeled bZiPro
21 at 25 °C. For the R_1 measurement, several 2D experiments with relaxation delays of 100, 200,
22 400, 600, 800, 1200, 1600, 2000, 2500 and 3000 ms were recorded and processed. For the R_2
23 measurement, 2D spectra with relaxation delays of 17, 34, 51, 68, 85, 102, 119, 136 and 153 ms
24 were collected and processed. The hetNOEs were obtained using two datasets that were collected
25 with and without initial proton saturation for a period of 3 s (33).

26 **Protein crystallization**

27 1 μL of protein at a concentration of 40 mg/ml was mixed with 1 μL of reservoir solution containing 0.1 M Sodium
28 acetate trihydrate pH 4.6, 0.2 M Ammonium sulfate, 30% PEG 2000 by hanging vapor diffusion method. Crystals
29 appeared at 18°C after two days. Crystals were flash-frozen in liquid N_2 before mounting. For crystallization with
30 EN300, bZiPro crystals were grown in 1 μL reservoir solution containing 2M Ammonium sulfate, 5% 2-propanol
31 with 1 μL of protein at 40 mg/ml at 18°C followed by soaking with 10 mM EN300 at 18°C for 4 hours. The
32 complex was cryoprotected in original crystallization solution with 20% glycerol and flash-cooled in liquid N_2 .

33 **Data collection, crystal structure determination and refinement**

34 For bZiPro X-ray diffraction data were acquired at the MX2 beamline at the Australian
35 Synchrotron (Clayton, Victoria). The complex diffraction data were collected at PXIII beamline
36 at Swiss Light Source (Paul Scherrer Institut, Villigen). Data collection was controlled using
37 Blu-Ice software. Diffraction intensities were integrated using iMOSFLM(34). Scaling and
38 merging of the intensities were carried out using software POINTLESS and AIMLESS from
39 CCP4 suite (35-37). Data collection statistics are summarized in **Table S1**. A solution for the
40 bZiPro crystal structure was determined by molecular replacement with the program MOLREP
41 (38) using eZiPro structure as a model (PDB code 5GJ4). The structure was subjected to
42 iterative rounds of refinement using the Phenix.refine program(39) and manual rebuilding using
43 Coot (40). Significant twinning was detected for free enzyme bZiPro crystals and the twin
44 fraction is 0.33. The twin law "k, h, -l" was applied during phenix refinement. Refinement
45 statistics are summarized in **Table S1**. Twinning was not observed in crystals with EN300.
46 Figures were generated using Pymol(41).

Thermal shift assay

The assay was conducted using the method described previously (33). Protein concentrations as 1mg/ml and the final concentration of SYPRO Orange as $5 \times$ were used for measurement. The dilution buffer was 25 mM HEPES pH 7.5, 75 mM NaCl, and 2 mM DTT. 10 % DMSO and 20 mM compound EN300 were added to the protein samples separately. The T_m values were calculated from the heat induced melting curves. Data was analyzed by Graph Pad Prism version 5.00 (La Jolla California, USA) using the normalization function.

Inhibition assay

Inhibition of Zika proteases by compound EN300 was measured as previous reported (24). Concentration of EN300 from 20 mM to 0.04 mM was incubated with 3 nM of bZiPro with 10% DMSO in buffer containing 20 mM Tris pH 8.5, 10% glycerol, 0.01% Triton X-100. The samples were incubated at room temperature for 60 minutes in a 96 well plate. The fluorescence intensities were measured after addition of Bz-nKRR-AMC substrate to a final concentration of 20 μ M by using Cytation 3 Multitmode plate reader (BioTek) with excitation wavelength as 380 nm and emission wavelength as 460 nm. The initial velocities were evaluated every 1 minute for over 10 minutes. Experiments were repeated 3 times at 37 °C. The data were analyzed using Transform function, Graph Pad. Samples from wells with enzyme without substrate nor EN300 were normalized as 0% and those from wells with enzyme and substrate without EN300 were normalized as 100%. The inhibition curve was plotted using linear regression function with X axis as Log₁₀ scale in Graph Pad.

Supplementary Figures

Figure S1. The refined structure consists of four molecules of bZiPro in one unit cell. (A) The four molecules are labeled as AB, CD, EF, and GH. There are three molecules of free enzyme (CD, EF, and GH) and one (AB) bound to a peptide ($K_{14}K_{15}G_{16}E_{17}$) from the neighbor NS3 N-terminus in an unusual reverse orientation. (B) Superposition of the four molecules. (C) Putty cartoon view of temperature factor variation on the crystal structure, colored from low to high values ($14\sim70\text{ \AA}^2$ as blue to red). (D) Simulated annealing mFo-DFc maps of the free enzymes are contoured at 3σ in green mesh. 2mFo-DFc electron density map is contoured at 1σ in blue. The figures are zoomed in at the substrate active site. (E) The representative electron density maps of the free enzymes after final refinement. Zoomed in at the same regions as in (D).

A

B

C

D

E

Figure S2. Crystal structures of the flaviviral NS2B-NS3 proteases. For DENV 1, 2, 3, JEV and WNV, NS3 from are colored in cyan and NS2B in green. For ZIKV, NS3 is colored in yellow and NS2B in magenta. Peptide inhibitors and BPTI are colored in red. All previously reported NS2B-NS3 protease crystal structures, including the most recent one from ZIKV (PDB codes 5LC0, 5T1V, and 5GXJ) used the single chain construct of NS2B47 and NS3 protease connected by a G4SG4 linker. For these free proteases in the first row, NS2B cofactor is in a partially structured open conformation. Upon inhibitor (peptide inhibitor or BPTI) binding, the C-terminal region of NS2B can fold into a beta-hairpin structure and interact with NS3 and directly participate in substrate / inhibitor binding (second row). We captured the bivalent ZIKV NS2B-NS3 protease (bottom right, free bZiPro and its complex with a small compound) in the closed conformation.

Figure S3. Structural views of eZiPro and gZiPro and the ligand binding interactions. (A) Electrostatic view and (C) interaction network of eZiPro in complex with the NS2B C-terminal peptide T₁₂₇G₁₂₈K₁₂₉R₁₃₀ post proteolysis. (B) Electrostatic view and (D) interaction network of gZiPro in complex with a peptidomimetic boronic acid inhibitor. Electrostatic surfaces are colored by electrostatic potential at neutral pH from −5 kT (red) to +5 kT (blue) using PyMOL.

Figure S4. A biological un-relevant of NS2B may present in both free and ligand bound bZiPro. (A) Overlay of ^1H - ^{15}N -HSQC spectra of bZiPro in the absence (black) and presence (red) of AcKR, and free NS2B cofactor region (green) is shown. No chemical shift perturbation was observed for the cross peaks overlapped with free NS2B cofactor region, suggesting that these peaks may be from free NS2B that dissociates with NS3. (B) Assignment of the ^1H - ^{15}N -HSQC spectrum of free bZiPro. (C) Assignment of the ^1H - ^{15}N -HSQC spectrum of bZiPro in complex with AcKR. The free NS2B is shown in green. Based on the peak intensities of G₈₂ of NS2B (free) and G₁₆ of NS3, approximately 25% free NS2B was estimated to be present in solution. Residues from NS2B that are biologically relevant are shown in red.

Figure S5. Assignment and secondary structure analysis of bZiPro. (A) The secondary structure of bZiPro was based on assigned residues using TALOSN (31, 42). The secondary structures of bZiPro in the absence (green) and presence of AcKR (red) were compared with those obtained from crystal structures of bZiPro free enzyme (Chains G/H). Residues having no resonance assigned in free bZiPro are highlighted in blue. For NS2B forming complex with NS3, residues 49-66 were assigned and their structures are similar to those of X-ray. Residues from 67-90 were broadened, which may be due to protein dynamics. (B) Secondary structure and dynamic analysis (C) of the form of NS2B that did not response to addition of AcKR. As there is no significant change in the both secondary structure and stability observed upon AcKR binding, suggesting that these peaks might come from a biological un-relevant of NS2B or free NS2B dissociated from NS3. The residues from the β hairpin region response to peptide are also highlighted in blue. Signals from these residues were broadened in free bZiPro.

Figure S6. Overlay of 2D ^1H - ^{15}N -HSQC spectra of eZiPro and bZiPro-AcKR. The spectra of eZiPro and bZiPro (0.8 mM) in the presence of 3.2 mM of AcKR were overlaid. Peaks from T₁₂₇ and G₁₂₈ of NS2B can be resolved as they are not present in bZiPro. G₁₅₃ exhibited chemical shift difference because of its close contact with T₁₂₇ and G₁₂₈ of NS2B in eZiPro. Residues such as H₅₁, G₁₃₃, A₁₃₂, and Y₁₆₁ in and close to the active triad, and residues such as K₁₁₇ and A₁₆₄ away from the active site exhibited line broadening in bZiPro-AcKR complex. This result indicates that varying the length of the peptide inhibitor (removing P3 and P4 amino acids) may affect the environment of the active site, which also affects the chemical environment or dynamic of residues away from the pocket.

Figure S7. (A)Residues affected by peptide binding. Chemical shift perturbation (CSP) of bZiPro upon AcKR peptide binding. CSP ($\Delta\delta$) were determined for assigned residues using the following equation: $\Delta\delta = ((\Delta\delta_{\text{HN}})^2 + (\Delta\delta_{\text{N}}/5)^2)^{0.5}$, where $\Delta\delta_{\text{HN}}$, and $\Delta\delta_{\text{N}}$ are the CSPs caused by AcKR binding in the amide proton and amide dimensions, respectively (43).

Figure S8. Relaxation analysis of bZiPro in the absence and presence of AcKR. The ^{15}N R_1 and hetNOE values of bZiPro in the absence and presence of AcKR peptide were plotted against residue number. The data were acquired using 0.8 mM of bZiPro the absence (black) and presence (red) 3.2 mM of AcKR. Protein was not stable at 37 °C for long-time experiments and the data were acquired at 25 °C. For NS2B cofactor, new peaks appeared for residues 66-87. These newly appeared peaks exhibited similar dynamics to structured residues of NS3, suggesting that they are forming a tight complex with NS3. No significant changes were observed for residues in both NS2B and NS3 in the absence and presence of AcKR, suggesting that free bZiPro forms a closed conformation in solution.

Figure S9. Stabilization and inhibition of bZiPro by compound EN300. (A) Fluorescence intensity changes of bZiPro and bZiPro with EN300. Melting temperatures for bZiPro alone is 46.14 °C and for bZiPro with EN300 is 46.82 °C. The T_m of the bZiPro exhibited a slightly shift (0.68 °C) in the presence of 20 mM EN300. (B) Heat

induced melting curves of bZiPro and bZiPro with EN300 complex. Thermostabilities of the bZiPro and bZiPro with EN300 complex were analyzed. (C) Inhibition of bZiPro activity by EN300. As shown below, IC_{50} was estimated to be 14.08 μM , indicating that EN300 negligibly inhibits the protease activity of bZiPro.

Table S1. Data Collection and Refinement Statistics

Data Collection Statistics	bZiPro	bZiPro+EN300
Wavelength (Å)	1.0428	1.0000
Resolution range (Å)	60.38 - 1.58 (1.60 - 1.58) *	43.00 - 2.00 (2.06 - 2.00)
Space group	P2 ₁ 2 ₁ 2 ₁	P4 ₃ 22
Unit cell a, b, c, α, β, γ (°)	60.27, 60.38, 215.46, 90.0, 90.0, 90.0	42.35, 42.35, 215.00 90.0, 90.0, 90.0
Total number of reflections	742298	135874
Unique reflections	108701	14167
Multiplicity	6.8	9.6
Completeness (%)	99.6 (95.4)	99.9 (99.5)
I/σI	7.4 (1.8)	14.6 (2.1)
Wilson B-factor (Å ²)	14.7	41.5
R _p	0.059 (0.309)	0.028 (0.295)
CC _{1/2}	0.987 (0.696)	0.999 (0.925)
^a R _{merge}	0.134 (0.647)	0.071 (0.859)
Refinement Statistics		
Resolution range (Å)	46.17-1.58 (1.63 - 1.58)	42.34 - 2.00 (2.08 - 2.00)
^b R _{work} (%)	15.98	21.15(30.50)
^c R _{free} (%)	19.65	24.43(30.34)
Number of non-hydrogen atoms	6331	1511
Macromolecules	5656	1430
Ligands	-	15
Water	656	66
Protein residues	761	193
RMSD (bonds) (Å)	0.005	0.004
^d RMSD (angles) (°)	1.13	1.04
Ramachandran favoured (%)	96.52	97.35
Ramachandran allowed (%)	3.35	2.65
Ramachandran outliers (%)	0.13	0.00
Clashscore	7.24	4.60
Average B-factor (Å ²)	26.0	55.6
Macromolecules	24.9	55.4
Ligands		74.5
Solvent	35.1	54.1

*The numbers in parentheses refer to the highest resolution shell.

^a R_{merge} = $\sum |I_j - \langle I \rangle| / \sum I_j$, where I_j is the intensity of an individual reflection, and $\langle I \rangle$ is the average intensity of that reflection.

^b R_{work} = $\sum ||F_o| - |F_c|| / \sum |F_c|$, where F_o denotes the observed structure factor amplitude, and F_c the structure factor amplitude calculated from the model.

^c R_{free} is as for R_{work} but calculated with 5% (3044) of randomly chosen reflections omitted from the refinement.

^dRMSD, root mean square deviations.

Table S2. Comparison of the crystal structures of ZIKV NS2B-NS3 proteases. Refer to **Fig S2**.

	bZiPro: Chains A/B	bZiPro: Chains C/D	bZiPro: Chains E/F	bZiPro: Chains G/H	bZiPro +EN300	eZiPro (5GJ4)	gZiPro (5LC0)
bZiPro: Chains A/B	-						
bZiPro: Chains C/D	0.31 Å /151 C α	-					
bZiPro: Chains E/F	0.26 Å /157 C α	0.22 Å /169 C α	-				
bZiPro: Chains G/H	0.34 Å /146 C α	0.17 Å /155 C α	0.24 Å /153 C α	-			
bZiPro +EN300	0.34 Å /164 C α	0.17 Å /155 C α	0.22 Å /160 C α	0.19 Å /154 C α	-		
eZiPro (5GJ4)	0.46 Å /159 C α	0.51 Å /165 C α	0.37 Å /171 C α	0.31 Å /164 C α	0.37 Å /166 C α	-	
gZiPro (5LC0)	0.58 Å /168 C α	0.51 Å /165 C α	0.52 Å /174 C α	0.52 Å /163 C α	0.51 Å /163 C α	0.53 Å /171 C α	-

1 **Author Contributions.**

2 Experiment design: CK and DL. Protein work for NMR experiments: YRL. Protein work, crystallization, data
3 collection, structure determination and analysis: ZZ, WP, YT and DL. NMR experiments: YL and CK. Fragment
4 characterization: ZZ and AWH. Manuscript writing: CK and DL with inputs from all authors.
5

References (23-43)

23. Y. M. Kim, S. Gayen, C. Kang, J. Joy, Q. Huang, A. S. Chen, J. L. Wee, M. J. Ang, H. A. Lim, A. W. Hung, R. Li, C. G. Noble, T. Lee le, A. Yip, Q. Y. Wang, C. S. Chia, J. Hill, P. Y. Shi, T. H. Keller, NMR analysis of a novel enzymatically active unlinked dengue NS2B-NS3 protease complex. *J Biol Chem* **288**, 12891-12900 (2013).
24. W. W. Phoo, Y. Li, Z. Zhang, M. Y. Lee, Y. R. Loh, Y. B. Tan, E. Y. Ng, J. Lescar, C. Kang, D. Luo, Structure of the NS2B-NS3 protease from Zika virus after self-cleavage. *Nature communications* **7**, 13410 (2016).
25. C. Nitsche, M. A. Behnam, C. Steuer, C. D. Klein, Retro peptide-hybrids as selective inhibitors of the Dengue virus NS2B-NS3 protease. *Antiviral Res* **94**, 72-79 (2012).
26. C. Kang, S. Gayen, W. Wang, R. Severin, A. S. Chen, H. A. Lim, C. S. Chia, A. Schuller, D. N. Doan, A. Poulsen, J. Hill, S. G. Vasudevan, T. H. Keller, Exploring the binding of peptidic West Nile virus NS2B-NS3 protease inhibitors by NMR. *Antiviral Res* **97**, 137-144 (2013).
27. K. Pervushin, A. Ono, C. Fernandez, T. Szyperski, M. Kainosho, K. Wuthrich, NMR scalar couplings across Watson-Crick base pair hydrogen bonds in DNA observed by transverse relaxation-optimized spectroscopy. *Proc Natl Acad Sci U S A* **95**, 14147-14151 (1998).
28. M. Salzmann, K. Pervushin, G. Wider, H. Senn, K. Wuthrich, TROSY in triple-resonance experiments: new perspectives for sequential NMR assignment of large proteins. *Proc Natl Acad Sci U S A* **95**, 13585-13590 (1998).
29. F. Delaglio, S. Grzesiek, G. W. Vuister, G. Zhu, J. Pfeifer, A. Bax, NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* **6**, 277-293 (1995).
30. B. A. Johnson, Using NMRView to visualize and analyze the NMR spectra of macromolecules. *Methods Mol Biol* **278**, 313-352 (2004).
31. Y. Shen, F. Delaglio, G. Cornilescu, A. Bax, TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J Biomol NMR* **44**, 213-223 (2009).
32. L. E. Kay, D. A. Torchia, A. Bax, Backbone dynamics of proteins as studied by ¹⁵N inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. *Biochemistry* **28**, 8972-8979 (1989).
33. S. Gayen, Q. Li, A. S. Chen, T. H. Nguyen, Q. Huang, J. Hill, C. Kang, An NMR study of the N-terminal domain of wild-type hERG and a T65P trafficking deficient hERG mutant. *Proteins* **79**, 2557-2565 (2011).
34. T. G. Battye, L. Kontogiannis, O. Johnson, H. R. Powell, A. G. Leslie, iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta crystallographica. Section D, Biological crystallography* **67**, 271-281 (2011).
35. P. R. Evans, An introduction to stereochemical restraints. *Acta crystallographica. Section D, Biological crystallography* **63**, 58-61 (2007).
36. P. R. Evans, An introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta crystallographica. Section D, Biological crystallography* **67**, 282-292 (2011).
37. E. Potterton, P. Briggs, M. Turkenburg, E. Dodson, A graphical user interface to the CCP4 program suite. *Acta crystallographica. Section D, Biological crystallography* **59**, 1131-1137 (2003).
38. A. Vagin, A. Teplyakov, MOLREP: an automated program for molecular replacement. *Journal of applied crystallography* **30**, 1022-1025 (1997).
39. P. D. Adams, P. V. Afonine, G. Bunkoczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, P. H. Zwart, PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-221 (2010).
40. P. Emsley, K. Cowtan, Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-2132 (2004).
41. L. Schrödinger, *The PyMOL Molecular Graphics System, Version 1.5.0.4* (2010).
42. Y. Shen, A. Bax, Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J Biomol NMR* **56**, 227-241 (2013).
43. M. P. Williamson, Using chemical shift perturbation to characterise ligand binding. *Progress in Nuclear Magnetic Resonance Spectroscopy* **73**, 1-16 (2013).