Identification and structural characterization of small molecule fragments targeting Zika virus NS2B-NS3 protease

Jun Ping Quek [1,#], Shuang Liu[2,4#], Zhenzhen Zhang[1] , Yan Li[2, 5], Elizabeth Yihui Ng[2], Ying Ru Loh[2], Alvin W. Hung[2,*], Dahai Luo[1,3,*], CongBao Kang [2,*]

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

[2]Experimental Drug Development Centre (EDDC), Agency for Science, Technology and Research (A*STAR), 10 Biopolis road, Chromos, #05-01, Singapore 138670.

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

Corresponding Authors

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

#These authors contributed equally.

[4] Shuang Liu's current address: Broad Institute of MIT and Harvard, 415 Main Street, Cambridge, MA 02142.

[5] Yan Li's present address: Department of Pathogen Biology, School of Basic Medicine, Tongji Medical College, Huazhong University of Science and Technology, 13 Hangkong Road, Wuhan, Hubei, P.R.China, 430030.

ABSTRACT

Zika virus (ZIKV) NS2B-NS3 protease is a validated antiviral target as it is essential for maturation of viral proteins. However, its negatively charged active site hinders the development of orthosteric small-molecule inhibitors. Fragment-based drug discovery (FBDD) is a powerful tool to generate novel chemical starting points against difficult drug targets. In this study, we screened a fragment compound library against the Zika protease using a primary thermal shift assay and identified twenty-two fragments which (bind to and) stabilize the protease. We then determined the X-ray crystal structures of two hits from different classes, all of which bind to the S1 pocket next to the protease active site. We confirmed that these two fragments bind to the protease without inducing significant conformational changes using solution NMR spectroscopy. These fragment scaffolds serve as the starting point for subsequent lead compound development

KEYWORDS: Zika Virus; Fragment-based drug discovery; protease inhibitor; structure; NMR; X-ray crystallography.

1 Introduction

Zika virus (ZIKV) is a mosquito-borne flavivirus whose family also encompasses other important human pathogens such as Dengue virus (DENV) and West Nile virus (WNV). ZIKV had been a neglected virus until its major outbreak in 2015, representing a global health threat due to its relationship to neurological disorders (Broutet et al., 2016; Ndeffo-Mbah et al., 2016). ZIKV is a small enveloped virus whose genome is a positive single-stranded RNA. The translated viral polyprotein is further processed into three structural proteins (capsid, membrane, and envelope) and seven nonstructural (NS) proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5) (Enfissi et al.; Kuno and Chang, 2007). ZIKV protease is a two-component complex formed by approximately 40 residues from the cytoplasmic region of NS2B and approximately 170 residues of NS3 (Lei et al., 2016). NS3 contains the catalytic triad (H51, D75 and S135) while residues from NS2B are critical for folding and enzymatic activity by forming a complex with NS3 and interacting with the substrate (Lei et al., 2016; Phoo et al., 2016). Overall, the protease cleaves the cytoplasmic-side joints of non-structural proteins required for viral assembly and replication (Uchil and Satchidanandam, 2003), making it an attractive antiviral drug target (Kang et al., 2017).

The folding of ZIKV protease is similar to other flavivirus proteases such as WNV and DENV (Lei et al., 2016; Li et al., 2017; Phoo et al., 2016; Zhang et al., 2016) and the active site is negatively charged (Li et al.; Phoo et al., 2016). Such an active site of the protease has a strong preference to positively-charged molecules for tight binding. However, the charged molecules have poor druggability characteristics due to the poor membrane permeability and oral absorption (Poulsen et al., 2014). As a result, the druggability of ZIKV protease is low and it is challenging to develop potent small molecular inhibitors using conventional approaches. Structural studies on the ZIKV protease in the absence and presence of different types of inhibitors have been carried out to provide valuable information to understand protease and ligand interactions (Chen et al., 2016; Lei et al., 2016; Li et al., 2017; Mahawaththa et al., 2017; Phoo et al., 2016; Phoo et al., 2018; Roy et al., 2017; Shiryaev et al., 2017; Zhang et al., 2016). Inhibitors derived from substrate have been developed with

nanomolar IC_{50}s against flavivirus proteases (Kang et al., 2017; Li et al.; Poulsen et al., 2014). Studies have also been carried out to explore allosteric inhibitors (Yao et al., 2019). Despite the progress in the development of a few small molecular weight compounds (Kumar et al., 2018; Shiryaev et al., 2017) and peptidic inhibitors (Nitsche et al., 2019), it remains challenging for these inhibitors to enter clinical studies due to potential issues such as cell penetration and stability in vivo. Therefore, small molecular weight compounds with different scaffolds and suitable for further development are desired to meet medical needs.

Fragment-based drug design (FBDD) is one popular drug discovery approach which starts by identifying small chemical fragments that bind to the biological target, and subsequently growing or linking the identified fragments to produce lead with higher binding affinity and potency (Erlanson et al., 2016). Compared to the conventional high-throughput screening, FBDD has several advantages such as quick access of structure-activity relationship, small size of the fragments covering a wide chemical space, higher ligand efficiency of the screened fragments, and only a small collection of fragments required for screening (Harner et al., 2013). This method has also been shown to work against more difficult targets such as protein-protein interactions and proteins with low druggabilities (Hajduk et al., 2005; Harner et al., 2013). A recent study shows that several fragments bind to different sites of WNV proteases, providing new chemical starting points for further drug development (Schöne et al., 2017). Previous, we reported the first fragment (EN300) which binds to the active site of the unlinked protease construct (bZiPro)(Lei et al., 2016; Li et al., 2017; Phoo et al., 2016; Zhang et al., 2016). Taken together, it is possible to design potent small molecular inhibitors against flavivirus proteases from structurally diverse fragments using FBDD approach.

In this study, we screened a fragment library using the eZiPro construct in which the cleavage product of the 2B-3 junction was present and the C-terminal residues of NS2B formed molecular interactions with the protease (Phoo et al., 2016). Twenty-two fragment compounds classified into four clusters were identified to increase the thermal stability of eZiPro by more than 0.6 ℃. X-ray

crystallographic studies demonstrate that two compounds bind to the active site of ZIKV protease. These two fragments bind to S1 pocket, which is similar to EN300 identified previously (Zhang et al., 2016). The molecular interactions between protease and two compounds were also confirmed using solution NMR spectroscopy. Although these fragments exhibited no detectable inhibitory activity on their own, our current study provides detailed information for newly identified fragments with different structures, which will be useful for growing them into potent inhibitors.

2 Materials and Methods

2.1 Protein production

Proteins including eZiPro and bZiPro (**Figure S1**) were produced as described previously (Phoo et al., 2016). Briefly, *E. coli* BL21(DE3) with plasmid was grown in LB or M9 medium. Protein was induced by adding IPTG when the cell density (OD600) reached 0.6-1.0. The cells were broken by sonication and the recombinant protein were purified using immobilized metal affinity chromatography and size-exclusion chromatography. Proteins were prepared in buffer A (20 mM HEPES pH 7.5, 150 mM NaCl, 5% glycerol, and 2 mM DTT) for x-ray crystallography and buffer B (20 mM HEPES, 150 mM and 2 mM DTT) for NMR and thermal shift assays.

2.2 Fragment screening

The screening was carried out against protease binding fragment using the eZiPro construct prepared in the buffer that contained 20 mM HEPES pH 7.3, 180 mM NaCl and 1 mM DTT. A compound library containing 1685 fragments was used in the screening. The assay was carried out in a white 384-well plate. In each well, the mixture contained 15 µM eZiPro, 20x SYPRO Orange dye and 2.5 mM of individual fragment. The samples were subjected to incremental temperature increases from 30 to 95 ℃ to obtain melting curves. The melting curves of eZiPro in the absence and presence of fragments were compared. Any fragment that can increase the stability of eZiPro 0.6 ℃ by (ΔTm ≥ 0.6 ℃) was considered as a positive hit. A hit rate of 1.3% was obtained for eZiPro, which was higher than that of bZiPro (0.2%) [6].

2.3 X-ray crystallographic studies

Compounds were prepared as a 100 mM stock solution by dissolving them in dimethyl sulfoxide. The bZiPro protease was used for crystallization which was described previously (Zhang et al., 2016). The compounds were added to protein (molar ratio 3:1) at a final protein concentration of 40 mg/ml for hanging vapor diffusion method. The bZiPro-compound complex was mixed with reservoir solution (0.2 M Ammonium sulfate, 0.1 M Sodium acetate trihydrate pH 4.6, 30% PEG2000) at 1:1 in volume ratio. A total of 2 µL was incubated at 18 oC until crystals appeared. Then the crystals were selected and soaked in cryoprotectant solution (0.2 M Ammonium sulfate, 0.1 M Sodium acetate trihydrate pH 4.6, 30% PEG2000, 20% glycerol), further flash froze in liquid N2. Diffraction data were collected and integrated using IMOSFLM Software (Battye et al., 2011). Data collection statistics were analyzed using POINTLESS and AIMLESS function of CCP4 suite (Evans, 2007, 2011; Potterton et al., 2003), as listed in Table 1. Refinement was processed by Phenix (Adams et al., 2010; Afonine et al., 2012; Headd et al., 2012) and Coot (Emsley and Cowtan, 2004; Emsley et al., 2010).

2.4 NMR experiments

Triple (^{13}C, ^{15}N and ^{2}H)-labeled bZiPro was prepared to a concentration of 0.6-0.8 mM in 20 mM HEPES, 150 mM, 90%H2O/10%D2O, and 2 mM DTT for acquiring ^{1}H-^{15}N-Heteronuclear single quantum coherence spectroscopy (HSQC). All the experiments were carried out 25 ℃ in a magnet with proton frequency of 600 MHz or 700 MHz. The ^{1}H-^{15}N-HSQC spectra of bZiPro in the absence and presence of compounds were acquired, processed and superimposed. The data were processed with NMRPipe (Delaglio et al., 1995) and visualized using NMRView (Johnson, 2004).

2.5 STD-NMR experiments

The saturation transfer difference NMR spectroscopy (STD-NMR) (Mayer and Meyer, 1999)was performed at 25 ℃ in a Bruker magnet with a proton frequency of 700 MHz equipped with a cryo-probe as described in the references (Schöne et al., 2017; Viegas et al., 2011) with slightly differences. The data were collected and processed using Bruker Topspin 2.1 software. The compounds were prepared in a 3 mm NMR tube. Each mixture containing 1 mM compound and 50 µM of bZiPro. Proton NMR was collected with solvent suppression. STD-NMR spectra were recorded

for each compound in the absence and presence of bZiPro using a pseudo-2D pulse sequence (stddiffesgp.3) (Topspin 2.1). The on- and off-frequencies were set at -111 Hz and 35000 Hz, respectively. The relaxation delay and saturation time were set at 3 s and 2 s, respectively. The protein signals were saturated using a selective Gaussian-shaped pulses (Gaus1.100) with a 50 ms duration.

3 Results

3.1 Identification of fragment compounds against eZiPro

In an effort to develop small molecule protease inhibitors, we embarked on a fragment screening approach to identify weak chemical hits that can be optimized into more potent leads. A low fragment hit rate was obtained when the bZiPro construct with an empty protease active site was used in screening (**Fig. 1**) (Zhang et al., 2016). We performed the same screen using a different protease construct-eZiPro in which the last four amino acids (TGKR) of NS2B form a complex with NS3 after cleavage and the interaction is undergoing fast exchange (Phoo et al., 2016). Thermal shift assays show the eZiPro (T_m= 50.9 ℃) construct being marginally more thermal stable than bZiPro (Tm= 48.9 ℃). As this construct is closer to the native form of viral protein and the TGKR sequence covers the protease active site and is flexible in solution (Figure S1), we then used it to screen fragment hits which might bind to multiple sites of the protease. Interestingly, a higher hit rate was observed for the eZiPro construct. More specifically, twenty-two fragments were identified to stabilize the eZiPro construct by more than 0.6 ℃ (**Fig. 1**, **Fig. S2**). The identified hits contain diverse chemical structures and were classified into four clusters based on the size and number of aromatic rings on fragments (**Fig. 1**). The first cluster consists of compounds with a 6 membered ring. The second cluster consists of fragments with a 5-member ring consisting of sulfur, carbon or nitrogen atoms. And the third and fourth clusters consist of two-ring structures, which are similar to those targeting WNV protease (Schöne et al., 2017). It is noteworthy that most hits from all three clusters contain a negatively carboxylic acid, suggesting electrostatic interactions between protease and

fragment hits. Next, we evaluated the inhibitory activities of these hit on the protease activity of ZIKV protease. Despite being detected for binding on the thermal shift, none of the fragments displayed measurable IC_{50} values (>200 μM). This is not surprising as the identified fragments are small, their binding to the protease is not strong enough to inhibit the protease activity. Further fragment growth is required to improve the potency.

3.2 Structures of bZiPro in complexes with fragments

To decipher the molecular interactions between the fragment hits and the protease, we co-crystallized these fragments with the bZiPro. We were able to determine the structures of bZiPro in complexes with two compounds, namely compounds **6** and **16** (**Fig. 2, Fig. S3 and Table S1**), at a resolution of 1.9 Å and 1.95 Å respectively. Interestingly, these two fragments bind to the same site on NS3 (P1 active site), similar to EN300-a fragment (Zhang et al., 2016). Although these two compounds bind to the same site on NS3, their interaction network are slightly different. Compound **6** has close contacts (<5 Å) with amino acids D129, Y130, P131, A132, S135, Y150, G151, and Y161. The hydroxyl group forms a hydrogen bond with the side chain oxygen of D129. Both oxygens in the carboxyl group can also participate to form hydrogen bonds with S135. The benzene ring in compound 6 is also capable of forming π-π interactions with Y161. Compound **16** has close contacts with amino acids including D129, Y130, A132, S135, Y150, G151 and Y161. Two possible hydrogen bonds are identified, namely the oxygen atom of compound 16 and the hydroxyl group of S135, and between the amide group of compound 16 and backbone oxygen of Y130.

Structures of NS2B and NS3 remain unchanged when bZiPro forms complex with the fragments. They are the same as those in eZiPro and bZiPro in the absence and presence of peptide and small molecular inhibitors (**Fig 2 and Fig. S4**). These two fragments occupy part of the S1 position at the active site, overlapping with the side chain of Arg amino acid at the P1 position of peptidic inhibitors and substrate (**Fig. S4**). Although the S1 position of the protease active site is negatively charged, it severs as a hot spot for fragment binding (**Fig. S4**). As the structures of these fragments are different

from that of the side chain of Arg, it might be possible introduce an unnatural amino acid containing identified fragment to replace the Arg residue in the P1 position in peptidic inhibitors. Such modifications might be able to improve potency and other properties such as cell-penetrating activity. As the S2 position is also negatively charged, growing the fragment to the S' site might be a feasible strategy as this site is more hydrophobic than S1 and S2. Overall, our structures provide detailed information for bZiPro in complexes with several fragments.

3.3 NMR studies on the identified fragments

To further explore the molecular interactions between proteases and compounds **6** and **16** , we carried out solution NMR studies. Superimposition of the ^{1}H-^{15}N-HSQC spectra of 0.6 mM bZiPro in the absence and presence of 2.4 mM compounds confirmed their molecular interactions in solution. Exchanges exist in the free flavivirus protease, giving rise to line broadening of signals for residues close the C-terminal β-hairpin region of NS2B and its neighboring residues from NS3 (Kim et al., 2013; Su et al., 2009; Zhang et al., 2016). Addition of inhibitors or substrate suppress the changes, which will cause the appearance of signals in the ^{1}H-^{15}N-HSQC of bZiPro. The modest chemical shift perturbations of the ^{1}H-^{15}N-HSQC spectrum of bZiPro induced by addition of compounds **6** and **16** suggested that the binding did not suppress the exchanges in the protease (**Fig. 3**).

Ligand-based STD-NMR studies have been used to identify fragments against WNV protease (Schöne et al., 2017). STD-NMR can be used for identifying ligands with diverse dissociation constants ranging from 10^{-8} molL^{-1} to 10^{-3} molL^{-1} and identifying a binding site of a compound when a reference molecular is available (Viegas et al., 2011). In addition to solution phase protein NMR, ligand-based STD NMR was used to test whether these two compounds can be used as reference or competitive binding molecules for the development and characterization of novel Zika protease inhibitors. All these compounds exhibited signals in the STD experiment in the presence of bZiPro. Compound **16** might not be suitable to serve as a reference molecule in STD experiment as it only has one set of proton signal at approximately 4 ppm which is close to the buffer signals and may

cause irradiation effect. In total, both STD and HSQC experiments confirm the molecular interactions between bZiPro and compounds **6** and **16** (**Fig. 4**).

4 Discussion

Different strategies such as high-throughput screening, substrate-based drug design, in-silico screening and FBDD have been used to develop DENV, WNV and ZIKV protease inhibitors (Koh-Stenta et al., 2015; Luo et al., 2015; Manzano et al., 2014; Mass et al., 2007; Nitsche et al., 2014; Noble and Shi, 2012). Due to the negatively charged active site, most potent flavivirus protease inhibitors were derived from the substrate (Kang et al., 2017). Removal or reducing the peptide characteristics is therefore necessary to make these developed inhibitors have a potency to be used in clinical studies. We have identified several fragments binding to the S1 pocket in NS3 (**Fig. 2**) (Zhang et al., 2016). As these fragments bind to the same site as those of substrates and peptidic inhibitors, the information obtained in this study will be useful for modifying the Arg residue in the peptidic inhibitors. In addition to making a cyclic peptide to improve the potency and stability, replacing the P1 Arg with an unnatural amino acid containing the identified fragments might be helpful for generating a drug-like molecule.

The structures obtained in current and previous studies show that the fragment compounds bind to the S1 position in the protease active site, suggesting that the S1 position is a hot spot for small molecule binding (**Fig. 2**). The fragments can be used as starting compounds for further growing into more potent inhibitors. It has been noted that challenges still remain in further development of these fragments as S2 position is negatively charged. Growing the fragments to the S1' position or introducing active chemical groups to generate irreversible compounds might be feasible strategies to obtain potent small molecular weight inhibitors (Gruba et al., 2018). Nonetheless, our structural studies provide detailed information for the protease and fragment complexes, providing guidance for developing inhibitors using FBDD approach.

Developing allosteric inhibitors is a strategy for targets whose active sites are not druggable. This strategy has been used for DENV protease and the mechanism of action of the compounds is to lock the inactive form of protease (Nitsche et al., 2019; Yao et al., 2019; Yildiz et al., 2013). Non-competitive ZIKV protease has also been developed [20]. Conventional biochemical assay is challenging to directly screen for allosteric inhibitors against ZIKV protease. FBDD is a powerful tool to develop non-competitive inhibitors when reference compounds are available (Skora and Jahnke, 2017). In this study, the two compounds were fully characterized and shown to interact with the protease active site. Together with EN300, these compounds can be used as spy molecules in fragment screening using other methods such as NMR to identify molecules binding to different sites.

In the current study, we used eZiPro protease construct mimicking the natural form of the protease under physiological conditions. We originally planned to screen some fragments that were able to bind to different sites other than the active site. If successful, these identified fragments would be useful for developing allosteric inhibitors. The identified twenty fragments can be classified into four clusters (**Fig. 1**). We were able to present the crystal structures of bZiPro in complexes with fragments representing these four clusters (**Fig. 2**) (Zhang et al., 2016). Unexpectedly, the fragments bind to S1 position at the active site. Using solution NMR spectroscopy, we confirmed the binding of these two fragments in solution. Using compound AcKR-weak bZiPro binding peptide as a reference, competition experiment suggested that the fragments bind to bZiPro with an affinity in μM to mM range. It is interesting to see that bZiPro and eZiPro exhibited different hit rates against the same compound library under similar screening conditions. The presence of the TGKR peptide in the eZiPro did not affect screening fragments binding to the active site. This may be due to the weak binding affinities (approximately 200 μM) between the TGKR peptide and protease (Li et al., 2018). Having a more thermally stable construct might be useful for increasing hit rate in FBDD. Based on the current study, a more potent inhibitor such as the in-reversible inhibitor that can form a tight complex is required for screening non-competitive fragments using thermal shift assay.

In conclusion, we identified over twenty fragment compounds targeting ZIKV NS2B-NS3 protease using thermal shift assay. Two compounds were characterized using X-ray crystallography and solution NMR spectroscopy. These characterized hits are useful for developing ZIKV protease inhibitors and serving as reference molecules to characterize other new developed inhibitors.

Author Contributions

A.H, D.L. and C.K. designed the experiment. Y.R.L. and E.Y.N prepared protein for NMR studies. S.L. and A.H. screened the fragments. Y.L. and C.K. conducted and analyzed NMR experiments. Q.J., Z.Z., and D.L. performed crystallization and structure determination. D.L. and C.K. wrote the paper with input from all authors. All authors reviewed and revised the manuscript. All authors have given approval to the final version of the manuscript.

ACKNOWLEDGMENTS

We thank scientists from Australian Light Source MX beam-line and Swiss Light Source PX beam-line for their help with diffraction data collection. We also appreciate Prof Ho Sup Yoon and Dr Hong Ye for the help of NMR experiments. This work was supported by (1) a start-up grant from Lee Kong Chian School of Medicine, Nanyang Technological University, (2) Singapore National Research Foundation grant NRF2016NRF-CRP001-063, (3) National Medical Research Council grant CBRG15May045, to D.L. (3) Singapore Ministry of Health's National Medical Research Council (NMRC/OFIRG/0051/2017) to C.K.

Conflict of interests

The authors declare no competing financial interests.

Figure Legends

Figure 1. Identified fragment compounds against bZiPro and eZiPro using thermal shift assay. A. Three fragments were identified against bZiPro. These fragments were screened in a previous study using bZiPro protease construct (Lei et al., 2016; Li et al., 2017; Phoo et al., 2016; Zhang et al., 2016). B. Twenty-two fragments were identified against eZiPro. The compounds that are able to shift the melting curve of eZiPro equal to or more than 0.6 ℃ are considered as positive hits. The chemical structures of the fragments and the temperatures shift differences are shown. The identified fragments are classified into four clusters (highlighted in different colors) based on the number of the rings.

Figure 2. Structures of bZiPro in complex with the identified fragments. Structure of bZiPro in complex with compound 6 and 16. The upper panel is the residues that have close contacts with the fragments. The NS2B and NS3 are shown as purple and yellow cartoons, respectively. The fragments are shown as orange, cyan and green sticks for compound 6 and 16 respectively. The interactions are shown in black dashed lines. The lower panel is the electron density maps (2mFo-Fc) of the fragments in the complex, coloured in blue and contoured at 1σ.

Figure 3. ^{1}H-^{15}N-HSQC spectra of bZiPro in the absence and presence of compound 6 (A) and 16 (B). The ^{1}H-^{15}N-HSQC spectra of bZiPro in the absence (black) and presence (red) of fragments are shown in black and red, respectively.

Figure 4. STD-NMR analysis bZiPro and fragments interactions. NMR spectra of compound 6 (A) and 16 (B recorded on a Bruker 700 MHz magnet. Spectrum at the upper panel is free proton NMR of the compound. Spectrum at the lower panel is the STD-NMR of the compound in the presence of bZiPro. The chemical structures of the fragments are shown. One peak exhibited highest intensity at the amide region is indicated with an asterisk. This peak might overlay with impurities as it was present when free compound 6 was used for recording the same STD experiment.

References

Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., McCoy, A.J., Moriarty, N.W., Oeffner, R., Read, R.J., Richardson, D.C., Richardson, J.S., Terwilliger, T.C., Zwart, P.H., 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta crystallographica. Section D, Biological crystallography 66, 213-221.

Afonine, P.V., Grosse-Kunstleve, R.W., Echols, N., Headd, J.J., Moriarty, N.W., Mustyakimov, M., Terwilliger, T.C., Urzhumtsev, A., Zwart, P.H., Adams, P.D., 2012. Towards automated crystallographic structure refinement with phenix.refine. Acta crystallographica. Section D, Biological crystallography 68, 352-367.

Battye, T.G., Kontogiannis, L., Johnson, O., Powell, H.R., Leslie, A.G., 2011. iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. Acta crystallographica. Section D, Biological crystallography 67, 271-281.

Broutet, N., Krauer, F., Riesen, M., Khalakdina, A., Almiron, M., Aldighieri, S., Espinal, M., Low, N., Dye, C., 2016. Zika Virus as a Cause of Neurologic Disorders. N Engl J Med 374, 1506-1509.

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

Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J., Bax, A., 1995. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR 6, 277-293.

Emsley, P., Cowtan, K., 2004. Coot: model-building tools for molecular graphics. Acta Crystallogr D 60, 2126-2132.

Emsley, P., Lohkamp, B., Scott, W.G., Cowtan, K., 2010. Features and development of Coot. Acta Crystallographica Section D 66, 486-501.

Enfissi, A., Codrington, J., Roosblad, J., Kazanji, M., Rousset, D., Zika virus genome from the Americas. The Lancet 387, 227-228.

Erlanson, D.A., Fesik, S.W., Hubbard, R.E., Jahnke, W., Jhoti, H., 2016. Twenty years on: the impact of fragments on drug discovery. Nat Rev Drug Discov 15, 605-619.

Evans, P.R., 2007. An introduction to stereochemical restraints. Acta crystallographica. Section D, Biological crystallography 63, 58-61.

Evans, P.R., 2011. An introduction to data reduction: space-group determination, scaling and intensity statistics. Acta crystallographica. Section D, Biological crystallography 67, 282-292.

Gruba, N., Martinez, J.I.R., Grzywa, R., Wysocka, M., Skorenski, M., Dabrowska, A., Lecka, M., Suder, P., Sienczyk, M., Pyrc, K., Lesner, A., 2018. One Step Beyond: Design of Substrates Spanning Primed Positions of Zika Virus NS2B-NS3 Protease. ACS Med Chem Lett 9, 1025-1029.

Hajduk, P.J., Huth, J.R., Fesik, S.W., 2005. Druggability indices for protein targets derived from NMR-based screening data. Journal of medicinal chemistry 48, 2518-2525.

Harner, M.J., Frank, A.O., Fesik, S.W., 2013. Fragment-based drug discovery using NMR spectroscopy. Journal of biomolecular NMR 56, 65-75.

Headd, J.J., Echols, N., Afonine, P.V., Grosse-Kunstleve, R.W., Chen, V.B., Moriarty, N.W., Richardson, D.C., Richardson, J.S., Adams, P.D., 2012. Use of knowledge-based restraints in phenix.refine to improve macromolecular refinement at low resolution. Acta crystallographica. Section D, Biological crystallography 68, 381-390.

Johnson, B.A., 2004. Using NMRView to visualize and analyze the NMR spectra of macromolecules. Methods Mol Biol 278, 313-352.

Kang, C., Keller, T.H., Luo, D., 2017. Zika Virus Protease: An Antiviral Drug Target. Trends in microbiology 25, 797-808.

Kim, Y.M., Gayen, S., Kang, C., Joy, J., Huang, Q., Chen, A.S., Wee, J.L., Ang, M.J., Lim, H.A., Hung, A.W., Li, R., Noble, C.G., Lee le, T., Yip, A., Wang, Q.Y., Chia, C.S., Hill, J., Shi, P.Y., Keller, T.H., 2013. NMR analysis of a novel enzymatically active unlinked dengue NS2B-NS3 protease complex. J Biol Chem 288, 12891-12900.
Koh-Stenta, X., Joy, J., Wang, S.F., Kwek, P.Z., Wee, J.L., Wan, K.F., Gayen, S., Chen, A.S., Kang, C., Lee, M.A., Poulsen, A., Vasudevan, S.G., Hill, J., Nacro, K., 2015. Identification of covalent active site inhibitors of dengue virus protease. Drug Des Devel Ther 9, 6389-6399.
Kumar, A., Liang, B., Aarthy, M., Singh, S.K., Garg, N., Mysorekar, I.U., Giri, R., 2018. Hydroxychloroquine Inhibits Zika Virus NS2B-NS3 Protease. ACS Omega 3, 18132-18141.
Kuno, G., Chang, J.G.-J., 2007. Full-length sequencing and genomic characterization of Bagaza, Kedougou, and Zika viruses. Archives of Virology 152, 687-696.
Lei, J., Hansen, G., Nitsche, C., Klein, C.D., Zhang, L., Hilgenfeld, R., 2016. Crystal structure of Zika virus NS2B-NS3 protease in complex with a boronate inhibitor. Science 353, 503-505.
Li, Y., Loh, Y.R., Hung, A.W., Kang, C., 2018. Characterization of molecular interactions between Zika virus protease and peptides derived from the C-terminus of NS2B. Biochem Biophys Res Commun 503, 691-696.
Li, Y., Phoo, W.W., Loh, Y.R., Zhang, Z., Ng, E.Y., Wang, W., Keller, T.H., Luo, D., Kang, C., 2017. Structural characterization of the linked NS2B-NS3 protease of Zika virus. FEBS Lett 591, 2338-2347.
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.
Luo, D., Vasudevan, S.G., Lescar, J., 2015. The flavivirus NS2B-NS3 protease-helicase as a target for antiviral drug development. Antiviral Res 118, 148-158.
Mahawaththa, M.C., Pearce, B.J., Szabo, M., Graham, B., Klein, C.D., Nitsche, C., Otting, G., 2017. Solution conformations of a linked construct of the Zika virus NS2B-NS3 protease. Antiviral Res 142, 141-147.
Manzano, M., Padia, J., Padmanabhan, R., 2014. Small molecule inhibitor discovery for dengue virus protease using high-throughput screening. Methods Mol Biol 1138, 331-344.
Mass, N., Selisko, B., Malet, H., Peyrane, F., Debarnot, C., Decroly, E., Benarroch, D., Egloff, M.P., Guillernot, J.C., Alvarez, K., Canard, B., 2007. Dengue virus: viral targets and antiviral drugs. Virologie (Montrouge) 11, 121-133.
Mayer, M., Meyer, B., 1999. Characterization of Ligand Binding by Saturation Transfer Difference NMR Spectroscopy. Angewandte Chemie International Edition 38, 1784-1788.
Ndeffo-Mbah, M.L., Parpia, A.S., Galvani, A.P., 2016. Mitigating Prenatal Zika Virus Infection in the Americas. Annals of internal medicine 165, 551-559.
Nitsche, C., Holloway, S., Schirmeister, T., Klein, C.D., 2014. Biochemistry and Medicinal Chemistry of the Dengue Virus Protease. Chemical Reviews.
Nitsche, C., Passioura, T., Varava, P., Mahawaththa, M.C., Leuthold, M.M., Klein, C.D., Suga, H., Otting, G., 2019. De Novo Discovery of Nonstandard Macrocyclic Peptides as Noncompetitive Inhibitors of the Zika Virus NS2B-NS3 Protease. ACS Med Chem Lett 10, 168-174.
Noble, C.G., Shi, P.Y., 2012. Structural biology of dengue virus enzymes: Towards rational design of therapeutics. Antiviral Res 96, 115-126.
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., 2016. Structure of the NS2B-NS3 protease from Zika virus after self-cleavage. Nat Commun 7, 13410.
Phoo, W.W., Zhang, Z., Wirawan, M., Chew, E.J.C., Chew, A.B.L., Kouretova, J., Steinmetzer, T., Luo, D., 2018. Structures of Zika virus NS2B-NS3 protease in complex with peptidomimetic inhibitors. Antiviral Research 160, 17-24.
Potterton, E., Briggs, P., Turkenburg, M., Dodson, E., 2003. A graphical user interface to the CCP4 program suite. Acta crystallographica. Section D, Biological crystallography 59, 1131-1137.
Poulsen, A., Kang, C., Keller, T.H., 2014. Drug design for flavivirus proteases: what are we missing? Curr Pharm Des 20, 3422-3427.

Roy, A., Lim, L., Srivastava, S., Lu, Y., Song, J., 2017. Solution conformations of Zika NS2B-NS3pro and its inhibition by natural products from edible plants. PloS one 12, e0180632.
Schöne, T., Grimm, L.L., Sakai, N., Zhang, L., Hilgenfeld, R., Peters, T., 2017. STD-NMR experiments identify a structural motif with novel second-site activity against West Nile virus NS2B-NS3 protease. Antiviral Research 146, 174-183.
Shiryaev, S.A., Farhy, C., Pinto, A., Huang, C.-T., Simonetti, N., Ngono, A.E., Dewing, A., Shresta, S., Pinkerton, A.B., Cieplak, P., Strongin, A.Y., Terskikh, A.V., 2017. Characterization of the Zika virus two-component NS2B-NS3 protease and structure-assisted identification of allosteric small-molecule antagonists. Antiviral Research 143, 218-229.
Skora, L., Jahnke, W., 2017. 19F-NMR-Based Dual-Site Reporter Assay for the Discovery and Distinction of Catalytic and Allosteric Kinase Inhibitors. ACS Medicinal Chemistry Letters 8, 632-635.
Su, X.C., Ozawa, K., Qi, R., Vasudevan, S.G., Lim, S.P., Otting, G., 2009. NMR analysis of the dynamic exchange of the NS2B cofactor between open and closed conformations of the West Nile virus NS2B-NS3 protease. PLoS Negl Trop Dis 3, e561.
Uchil, P.D., Satchidanandam, V., 2003. Architecture of the Flaviviral Replication Complex: PROTEASE, NUCLEASE, AND DETERGENTS REVEAL ENCASEMENT WITHIN DOUBLE-LAYERED MEMBRANE COMPARTMENTS. Journal of Biological Chemistry 278, 24388-24398.
Viegas, A., Manso, J., Nobrega, F.L., Cabrita, E.J., 2011. Saturation-Transfer Difference (STD) NMR: A Simple and Fast Method for Ligand Screening and Characterization of Protein Binding. Journal of Chemical Education 88, 990-994.
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., 2019. Discovery, X-ray Crystallography and Antiviral Activity of Allosteric Inhibitors of Flavivirus NS2B-NS3 Protease. J Am Chem Soc 141, 6832-6836.
Yildiz, M., Ghosh, S., Bell, J.A., Sherman, W., Hardy, J.A., 2013. Allosteric Inhibition of the NS2B-NS3 Protease from Dengue Virus. ACS Chem Biol 17, 76-80.
Zhang, Z., Li, Y., Loh, Y.R., Phoo, W.W., Hung, A.W., Kang, C., Luo, D., 2016. Crystal structure of unlinked NS2B-NS3 protease from Zika virus. Science 354, 1597-1600.

Fragment screening was carried out against Zika virus protease.

Co-crystal structures of the identified hits with proteases were obtained.

S1 is a hot spot for fragment binding.

Fragment binding to protease does not induce significant structural changes.

Identification and structural characterization of small molecule fragments targeting Zika virus NS2B-NS3 protease

Jun Ping Quek [1,#], Shuang Liu[2,4#], Zhenzhen Zhang[1] , Yan Li[2, 5], Elizabeth Yihui Ng[2], Ying Ru Loh[2], Alvin W. Hung[2,*], Dahai Luo[1,3,*], CongBao Kang [2,*]

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

[2]Experimental Drug Development Centre (EDDC), Agency for Science, Technology and Research (A*STAR), 10 Biopolis road, Chromos, #05-01, Singapore 138670.

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

Corresponding Authors

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

#These authors contributed equally.

[4] Shuang Liu's current address: Broad Institute of MIT and Harvard, 415 Main Street, Cambridge, MA 02142.

[5] Yan Li's present address: Department of Pathogen Biology, School of Basic Medicine, Tongji Medical College, Huazhong University of Science and Technology, 13 Hangkong Road, Wuhan, Hubei, P.R.China, 430030.

ABSTRACT

Zika virus (ZIKV) NS2B-NS3 protease is a validated antiviral target as it is essential for maturation of viral proteins. However, its negatively charged active site hinders the development of orthosteric small-molecule inhibitors. Fragment-based drug discovery (FBDD) is a powerful tool to generate novel chemical starting points against difficult drug targets. In this study, we screened a fragment compound library against the Zika protease using a primary thermal shift assay and identified twenty-two fragments which (bind to and) stabilize the protease. We then determined the X-ray crystal structures of two hits from different classes, all of which bind to the S1 pocket next to the protease active site. We confirmed that these two fragments bind to the protease without inducing significant conformational changes using solution NMR spectroscopy. These fragment scaffolds serve as the starting point for subsequent lead compound development

KEYWORDS: Zika Virus; Fragment-based drug discovery; protease inhibitor; structure; NMR; X-ray crystallography.

1 Introduction

Zika virus (ZIKV) is a mosquito-borne flavivirus whose family also encompasses other important human pathogens such as Dengue virus (DENV) and West Nile virus (WNV). ZIKV had been a neglected virus until its major outbreak in 2015, representing a global health threat due to its relationship to neurological disorders (Broutet et al., 2016; Ndeffo-Mbah et al., 2016). ZIKV is a small enveloped virus whose genome is a positive single-stranded RNA. The translated viral polyprotein is further processed into three structural proteins (capsid, membrane, and envelope) and seven nonstructural (NS) proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5) (Enfissi et al.; Kuno and Chang, 2007). ZIKV protease is a two-component complex formed by approximately 40 residues from the cytoplasmic region of NS2B and approximately 170 residues of NS3 (Lei et al., 2016). NS3 contains the catalytic triad (H51, D75 and S135) while residues from NS2B are critical for folding and enzymatic activity by forming a complex with NS3 and interacting with the substrate (Lei et al., 2016; Phoo et al., 2016). Overall, the protease cleaves the cytoplasmic-side joints of non-structural proteins required for viral assembly and replication (Uchil and Satchidanandam, 2003), making it an attractive antiviral drug target (Kang et al., 2017).

The folding of ZIKV protease is similar to other flavivirus proteases such as WNV and DENV (Lei et al., 2016; Li et al., 2017; Phoo et al., 2016; Zhang et al., 2016) and the active site is negatively charged (Li et al.; Phoo et al., 2016). Such an active site of the protease has a strong preference to positively-charged molecules for tight binding. However, the charged molecules have poor druggability characteristics due to the poor membrane permeability and oral absorption (Poulsen et al., 2014). As a result, the druggability of ZIKV protease is low and it is challenging to develop potent small molecular inhibitors using conventional approaches. Structural studies on the ZIKV protease in the absence and presence of different types of inhibitors have been carried out to provide valuable information to understand protease and ligand interactions (Chen et al., 2016; Lei et al., 2016; Li et al., 2017; Mahawaththa et al., 2017; Phoo et al., 2016; Phoo et al., 2018; Roy et al., 2017; Shiryaev et al., 2017; Zhang et al., 2016). Inhibitors derived from substrate have been developed with nanomolar IC_{50}s

against flavivirus proteases (Kang et al., 2017; Li et al.; Poulsen et al., 2014). Studies have also been carried out to explore allosteric inhibitors (Yao et al., 2019). Despite the progress in the development of a few small molecular weight compounds (Kumar et al., 2018; Shiryaev et al., 2017) and peptidic inhibitors (Nitsche et al., 2019), it remains challenging for these inhibitors to enter clinical studies due to potential issues such as cell penetration and stability in vivo. Therefore, small molecular weight compounds with different scaffolds and suitable for further development are desired to meet medical needs.

Fragment-based drug design (FBDD) is one popular drug discovery approach which starts by identifying small chemical fragments that bind to the biological target, and subsequently growing or linking the identified fragments to produce lead with higher binding affinity and potency (Erlanson et al., 2016). Compared to the conventional high-throughput screening, FBDD has several advantages such as quick access of structure-activity relationship, small size of the fragments covering a wide chemical space, higher ligand efficiency of the screened fragments, and only a small collection of fragments required for screening (Harner et al., 2013). This method has also been shown to work against more difficult targets such as protein-protein interactions and proteins with low druggabilities (Hajduk et al., 2005; Harner et al., 2013). A recent study shows that several fragments bind to different sites of WNV proteases, providing new chemical starting points for further drug development (Schöne et al., 2017). Previous, we reported the first fragment (EN300) which binds to the active site of the unlinked protease construct (bZiPro)(Lei et al., 2016; Li et al., 2017; Phoo et al., 2016; Zhang et al., 2016). Taken together, it is possible to design potent small molecular inhibitors against flavivirus proteases from structurally diverse fragments using FBDD approach.

In this study, we screened a fragment library using the eZiPro construct in which the cleavage product of the 2B-3 junction was present and the C-terminal residues of NS2B formed molecular interactions with the protease (Phoo et al., 2016). Twenty-two fragment compounds classified into four clusters were identified to increase the thermal stability of eZiPro by more than 0.6 ℃. X-ray crystallographic

studies demonstrate that two compounds bind to the active site of ZIKV protease. These two fragments bind to S1 pocket, which is similar to EN300 identified previously (Zhang et al., 2016). The molecular interactions between protease and two compounds were also confirmed using solution NMR spectroscopy. Although these fragments exhibited no detectable inhibitory activity on their own, our current study provides detailed information for newly identified fragments with different structures, which will be useful for growing them into potent inhibitors.

2 Materials and Methods

2.1 Protein production

Proteins including eZiPro and bZiPro (**Figure S1**) were produced as described previously (Phoo et al., 2016). Briefly, *E. coli* BL21(DE3) with plasmid was grown in LB or M9 medium. Protein was induced by adding IPTG when the cell density (OD600) reached 0.6-1.0. The cells were broken by sonication and the recombinant protein were purified using immobilized metal affinity chromatography and size-exclusion chromatography. Proteins were prepared in buffer A (20 mM HEPES pH 7.5, 150 mM NaCl, 5% glycerol, and 2 mM DTT) for x-ray crystallography and buffer B (20 mM HEPES, 150 mM and 2 mM DTT) for NMR and thermal shift assays.

2.2 Fragment screening

The screening was carried out against protease binding fragment using the eZiPro construct prepared in the buffer that contained 20 mM HEPES pH 7.3, 180 mM NaCl and 1 mM DTT. A compound library containing 1685 fragments was used in the screening. The assay was carried out in a white 384-well plate. In each well, the mixture contained 15 µM eZiPro, 20x SYPRO Orange dye and 2.5 mM of individual fragment. The samples were subjected to incremental temperature increases from 30 to 95 ℃ to obtain melting curves. The melting curves of eZiPro in the absence and presence of fragments were compared. Any fragment that can increase the stability of eZiPro 0.6 ℃ by (ΔTm ≥ 0.6 ℃) was considered as a positive hit. A hit rate of 1.3% was obtained for eZiPro, which was higher than that of bZiPro (0.2%) [6].

2.3 X-ray crystallographic studies

Compounds were prepared as a 100 mM stock solution by dissolving them in dimethyl sulfoxide. The bZiPro protease was used for crystallization which was described previously (Zhang et al., 2016). The compounds were added to protein (molar ratio 3:1) at a final protein concentration of 40 mg/ml for hanging vapor diffusion method. The bZiPro-compound complex was mixed with reservoir solution (0.2 M Ammonium sulfate, 0.1 M Sodium acetate trihydrate pH 4.6, 30% PEG2000) at 1:1 in volume ratio. A total of 2 µL was incubated at 18 oC until crystals appeared. Then the crystals were selected and soaked in cryoprotectant solution (0.2 M Ammonium sulfate, 0.1 M Sodium acetate trihydrate pH 4.6, 30% PEG2000, 20% glycerol), further flash froze in liquid N2. Diffraction data were collected and integrated using IMOSFLM Software (Battye et al., 2011). Data collection statistics were analyzed using POINTLESS and AIMLESS function of CCP4 suite (Evans, 2007, 2011; Potterton et al., 2003), as listed in Table 1. Refinement was processed by Phenix (Adams et al., 2010; Afonine et al., 2012; Headd et al., 2012) and Coot (Emsley and Cowtan, 2004; Emsley et al., 2010).

2.4 NMR experiments

Triple (^{13}C, ^{15}N and ^{2}H)-labeled bZiPro was prepared to a concentration of 0.6-0.8 mM in 20 mM HEPES, 150 mM, 90%H2O/10%D2O, and 2 mM DTT for acquiring ^{1}H-^{15}N-Heteronuclear single quantum coherence spectroscopy (HSQC). All the experiments were carried out 25 ℃ in a magnet with proton frequency of 600 MHz or 700 MHz. The ^{1}H-^{15}N-HSQC spectra of bZiPro in the absence and presence of compounds were acquired, processed and superimposed. The data were processed with NMRPipe (Delaglio et al., 1995) and visualized using NMRView (Johnson, 2004).

2.5 STD-NMR experiments

The saturation transfer difference NMR spectroscopy (STD-NMR) (Mayer and Meyer, 1999)was performed at 25 ℃ in a Bruker magnet with a proton frequency of 700 MHz equipped with a cryo-probe as described in the references (Schöne et al., 2017; Viegas et al., 2011) with slightly differences. The data were collected and processed using Bruker Topspin 2.1 software. The compounds were prepared in a 3 mm NMR tube. Each mixture containing 1 mM compound and 50 µM of bZiPro. Proton NMR was collected with solvent suppression. STD-NMR spectra were recorded for each compound in

the absence and presence of bZiPro using a pseudo-2D pulse sequence (stddiffesgp.3) (Topspin 2.1). The on- and off-frequencies were set at -111 Hz and 35000 Hz, respectively. The relaxation delay and saturation time were set at 3 s and 2 s, respectively. The protein signals were saturated using a selective Gaussian-shaped pulses (Gaus1.100) with a 50 ms duration.

3 Results

3.1 Identification of fragment compounds against eZiPro

In an effort to develop small molecule protease inhibitors, we embarked on a fragment screening approach to identify weak chemical hits that can be optimized into more potent leads. A low fragment hit rate was obtained when the bZiPro construct with an empty protease active site was used in screening (**Fig. 1**) (Zhang et al., 2016). We performed the same screen using a different protease construct-eZiPro in which the last four amino acids (TGKR) of NS2B form a complex with NS3 after cleavage and the interaction is undergoing fast exchange (Phoo et al., 2016). Thermal shift assays show the eZiPro (T_m= 50.9 ℃) construct being marginally more thermal stable than bZiPro (Tm= 48.9 ℃). As this construct is closer to the native form of viral protein and the TGKR sequence covers the protease active site and is flexible in solution (Figure S1), we then used it to screen fragment hits which might bind to multiple sites of the protease. Interestingly, a higher hit rate was observed for the eZiPro construct. More specifically, twenty-two fragments were identified to stabilize the eZiPro construct by more than 0.6 ℃ (**Fig. 1**, **Fig. S2**). The identified hits contain diverse chemical structures and were classified into four clusters based on the size and number of aromatic rings on fragments (**Fig. 1**). The first cluster consists of compounds with a 6 membered ring. The second cluster consists of fragments with a 5-member ring consisting of sulfur, carbon or nitrogen atoms. And the third and fourth clusters consist of two-ring structures, which are similar to those targeting WNV protease (Schöne et al., 2017). It is noteworthy that most hits from all three clusters contain a negatively carboxylic acid, suggesting electrostatic interactions between protease and fragment hits. Next, we evaluated the inhibitory activities of these hit on the protease activity of ZIKV protease. Despite being detected for binding on

the thermal shift, none of the fragments displayed measurable IC_{50} values (>200 μM). This is not surprising as the identified fragments are small, their binding to the protease is not strong enough to inhibit the protease activity. Further fragment growth is required to improve the potency.

3.2 Structures of bZiPro in complexes with fragments

To decipher the molecular interactions between the fragment hits and the protease, we co-crystallized these fragments with the bZiPro. We were able to determine the structures of bZiPro in complexes with two compounds, namely compounds **6** and **16** (**Fig. 2, Fig. S3 and Table S1**), at a resolution of 1.9 Å and 1.95 Å respectively. Interestingly, these two fragments bind to the same site on NS3 (P1 active site), similar to EN300-a fragment (Zhang et al., 2016). Although these two compounds bind to the same site on NS3, their interaction network are slightly different. Compound **6** has close contacts (<5 Å) with amino acids D129, Y130, P131, A132, S135, Y150, G151, and Y161. The hydroxyl group forms a hydrogen bond with the side chain oxygen of D129. Both oxygens in the carboxyl group can also participate to form hydrogen bonds with S135. The benzene ring in compound 6 is also capable of forming π-π interactions with Y161. Compound **16** has close contacts with amino acids including D129, Y130, A132, S135, Y150, G151 and Y161. Two possible hydrogen bonds are identified, namely the oxygen atom of compound 16 and the hydroxyl group of S135, and between the amide group of compound 16 and backbone oxygen of Y130.

Structures of NS2B and NS3 remain unchanged when bZiPro forms complex with the fragments. They are the same as those in eZiPro and bZiPro in the absence and presence of peptide and small molecular inhibitors (**Fig 2 and Fig. S4**). These two fragments occupy part of the S1 position at the active site, overlapping with the side chain of Arg amino acid at the P1 position of peptidic inhibitors and substrate (**Fig. S4**). Although the S1 position of the protease active site is negatively charged, it severs as a hot spot for fragment binding (**Fig. S4**). As the structures of these fragments are different from that of the side chain of Arg, it might be possible introduce an unnatural amino acid containing identified fragment to replace the Arg residue in the P1 position in peptidic inhibitors. Such modifications might

be able to improve potency and other properties such as cell-penetrating activity. As the S2 position is also negatively charged, growing the fragment to the S' site might be a feasible strategy as this site is more hydrophobic than S1 and S2. Overall, our structures provide detailed information for bZiPro in complexes with several fragments.

3.3 NMR studies on the identified fragments

To further explore the molecular interactions between proteases and compounds **6** and **16** , we carried out solution NMR studies. Superimposition of the ^{1}H-^{15}N-HSQC spectra of 0.6 mM bZiPro in the absence and presence of 2.4 mM compounds confirmed their molecular interactions in solution. Exchanges exist in the free flavivirus protease, giving rise to line broadening of signals for residues close the C-terminal β-hairpin region of NS2B and its neighboring residues from NS3 (Kim et al., 2013; Su et al., 2009; Zhang et al., 2016). Addition of inhibitors or substrate suppress the changes, which will cause the appearance of signals in the ^{1}H-^{15}N-HSQC of bZiPro. The modest chemical shift perturbations of the ^{1}H-^{15}N-HSQC spectrum of bZiPro induced by addition of compounds **6** and **16** suggested that the binding did not suppress the exchanges in the protease (**Fig. 3**).

Ligand-based STD-NMR studies have been used to identify fragments against WNV protease (Schöne et al., 2017). STD-NMR can be used for identifying ligands with diverse dissociation constants ranging from 10^{-8} molL^{-1} to 10^{-3} molL^{-1} and identifying a binding site of a compound when a reference molecular is available (Viegas et al., 2011). In addition to solution phase protein NMR, ligand-based STD NMR was used to test whether these two compounds can be used as reference or competitive binding molecules for the development and characterization of novel Zika protease inhibitors. All these compounds exhibited signals in the STD experiment in the presence of bZiPro. Compound **16** might not be suitable to serve as a reference molecule in STD experiment as it only has one set of proton signal at approximately 4 ppm which is close to the buffer signals and may cause irradiation effect. In total, both STD and HSQC experiments confirm the molecular interactions between bZiPro and compounds **6** and **16** (**Fig. 4**).

4 Discussion

Different strategies such as high-throughput screening, substrate-based drug design, in-silico screening and FBDD have been used to develop DENV, WNV and ZIKV protease inhibitors (Koh-Stenta et al., 2015; Luo et al., 2015; Manzano et al., 2014; Mass et al., 2007; Nitsche et al., 2014; Noble and Shi, 2012). Due to the negatively charged active site, most potent flavivirus protease inhibitors were derived from the substrate (Kang et al., 2017). Removal or reducing the peptide characteristics is therefore necessary to make these developed inhibitors have a potency to be used in clinical studies. We have identified several fragments binding to the S1 pocket in NS3 (**Fig. 2**) (Zhang et al., 2016). As these fragments bind to the same site as those of substrates and peptidic inhibitors, the information obtained in this study will be useful for modifying the Arg residue in the peptidic inhibitors. In addition to making a cyclic peptide to improve the potency and stability, replacing the P1 Arg with an unnatural amino acid containing the identified fragments might be helpful for generating a drug-like molecule.

The structures obtained in current and previous studies show that the fragment compounds bind to the S1 position in the protease active site, suggesting that the S1 position is a hot spot for small molecule binding (**Fig. 2**). The fragments can be used as starting compounds for further growing into more potent inhibitors. It has been noted that challenges still remain in further development of these fragments as S2 position is negatively charged. Growing the fragments to the S1' position or introducing active chemical groups to generate irreversible compounds might be feasible strategies to obtain potent small molecular weight inhibitors (Gruba et al., 2018). Nonetheless, our structural studies provide detailed information for the protease and fragment complexes, providing guidance for developing inhibitors using FBDD approach.

Developing allosteric inhibitors is a strategy for targets whose active sites are not druggable. This strategy has been used for DENV protease and the mechanism of action of the compounds is to lock the inactive form of protease (Nitsche et al., 2019; Yao et al., 2019; Yildiz et al., 2013). Non-competitive ZIKV protease has also been developed [20]. Conventional biochemical assay is challenging to directly

screen for allosteric inhibitors against ZIKV protease. FBDD is a powerful tool to develop non-competitive inhibitors when reference compounds are available (Skora and Jahnke, 2017). In this study, the two compounds were fully characterized and shown to interact with the protease active site. Together with EN300, these compounds can be used as spy molecules in fragment screening using other methods such as NMR to identify molecules binding to different sites.

In the current study, we used eZiPro protease construct mimicking the natural form of the protease under physiological conditions. We originally planned to screen some fragments that were able to bind to different sites other than the active site. If successful, these identified fragments would be useful for developing allosteric inhibitors. The identified twenty fragments can be classified into four clusters (**Fig. 1**). We were able to present the crystal structures of bZiPro in complexes with fragments representing these four clusters (**Fig. 2**) (Zhang et al., 2016). Unexpectedly, the fragments bind to S1 position at the active site. Using solution NMR spectroscopy, we confirmed the binding of these two fragments in solution. Using compound AcKR-weak bZiPro binding peptide as a reference, competition experiment suggested that the fragments bind to bZiPro with an affinity in μM to mM range. It is interesting to see that bZiPro and eZiPro exhibited different hit rates against the same compound library under similar screening conditions. The presence of the TGKR peptide in the eZiPro did not affect screening fragments binding to the active site. This may be due to the weak binding affinities (approximately 200 μM) between the TGKR peptide and protease (Li et al., 2018). Having a more thermally stable construct might be useful for increasing hit rate in FBDD. Based on the current study, a more potent inhibitor such as the in-reversible inhibitor that can form a tight complex is required for screening non-competitive fragments using thermal shift assay.

In conclusion, we identified over twenty fragment compounds targeting ZIKV NS2B-NS3 protease using thermal shift assay. Two compounds were characterized using X-ray crystallography and solution NMR spectroscopy. These characterized hits are useful for developing ZIKV protease inhibitors and serving as reference molecules to characterize other new developed inhibitors.

Author Contributions

A.H, D.L. and C.K. designed the experiment. Y.R.L. and E.Y.N prepared protein for NMR studies. S.L. and A.H. screened the fragments. Y.L. and C.K. conducted and analyzed NMR experiments. Q.J., Z.Z., and D.L. performed crystallization and structure determination. D.L. and C.K. wrote the paper with input from all authors. All authors reviewed and revised the manuscript. All authors have given approval to the final version of the manuscript.

ACKNOWLEDGMENTS

We thank scientists from Australian Light Source MX beam-line and Swiss Light Source PX beam-line for their help with diffraction data collection. We also appreciate Prof Ho Sup Yoon and Dr Hong Ye for the help of NMR experiments. This work was supported by (1) a start-up grant from Lee Kong Chian School of Medicine, Nanyang Technological University, (2) Singapore National Research Foundation grant NRF2016NRF-CRP001-063, (3) National Medical Research Council grant CBRG15May045, to D.L. (3) Singapore Ministry of Health's National Medical Research Council (NMRC/OFIRG/0051/2017) to C.K.

Conflict of interests

The authors declare no competing financial interests.

Figure Legends

Figure 1. Identified fragment compounds against bZiPro and eZiPro using thermal shift assay. A. Three fragments were identified against bZiPro. These fragments were screened in a previous study using bZiPro protease construct (Lei et al., 2016; Li et al., 2017; Phoo et al., 2016; Zhang et al., 2016). B. Twenty-two fragments were identified against eZiPro. The compounds that are able to shift the melting curve of eZiPro equal to or more than 0.6 ℃ are considered as positive hits. The chemical structures of the fragments and the temperatures shift differences are shown. The identified fragments are classified into four clusters (highlighted in different colors) based on the number of the rings.

Figure 2. Structures of bZiPro in complex with the identified fragments. Structure of bZiPro in complex with compound 6 and 16. The upper panel is the residues that have close contacts with the fragments. The NS2B and NS3 are shown as purple and yellow cartoons, respectively. The fragments are shown as orange, cyan and green sticks for compound 6 and 16 respectively. The interactions are shown in black dashed lines. The lower panel is the electron density maps (2mFo-Fc) of the fragments in the complex, coloured in blue and contoured at 1σ.

Figure 3. ^{1}H-^{15}N-HSQC spectra of bZiPro in the absence and presence of compound 6 (A) and 16 (B). The ^{1}H-^{15}N-HSQC spectra of bZiPro in the absence (black) and presence (red) of fragments are shown in black and red, respectively.

Figure 4. STD-NMR analysis bZiPro and fragments interactions. NMR spectra of compound 6 (A) and 16 (B recorded on a Bruker 700 MHz magnet. Spectrum at the upper panel is free proton NMR of the compound. Spectrum at the lower panel is the STD-NMR of the compound in the presence of bZiPro. The chemical structures of the fragments are shown. One peak exhibited highest intensity at the amide region is indicated with an asterisk. This peak might overlay with impurities as it was present when free compound 6 was used for recording the same STD experiment.

References

Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., McCoy, A.J., Moriarty, N.W., Oeffner, R., Read, R.J., Richardson, D.C., Richardson, J.S., Terwilliger, T.C., Zwart, P.H., 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta crystallographica. Section D, Biological crystallography 66, 213-221.

Afonine, P.V., Grosse-Kunstleve, R.W., Echols, N., Headd, J.J., Moriarty, N.W., Mustyakimov, M., Terwilliger, T.C., Urzhumtsev, A., Zwart, P.H., Adams, P.D., 2012. Towards automated crystallographic structure refinement with phenix.refine. Acta crystallographica. Section D, Biological crystallography 68, 352-367.

Battye, T.G., Kontogiannis, L., Johnson, O., Powell, H.R., Leslie, A.G., 2011. iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. Acta crystallographica. Section D, Biological crystallography 67, 271-281.

Broutet, N., Krauer, F., Riesen, M., Khalakdina, A., Almiron, M., Aldighieri, S., Espinal, M., Low, N., Dye, C., 2016. Zika Virus as a Cause of Neurologic Disorders. N Engl J Med 374, 1506-1509.

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

Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J., Bax, A., 1995. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR 6, 277-293.

Emsley, P., Cowtan, K., 2004. Coot: model-building tools for molecular graphics. Acta Crystallogr D 60, 2126-2132.

Emsley, P., Lohkamp, B., Scott, W.G., Cowtan, K., 2010. Features and development of Coot. Acta Crystallographica Section D 66, 486-501.

Enfissi, A., Codrington, J., Roosblad, J., Kazanji, M., Rousset, D., Zika virus genome from the Americas. The Lancet 387, 227-228.

Erlanson, D.A., Fesik, S.W., Hubbard, R.E., Jahnke, W., Jhoti, H., 2016. Twenty years on: the impact of fragments on drug discovery. Nat Rev Drug Discov 15, 605-619.

Evans, P.R., 2007. An introduction to stereochemical restraints. Acta crystallographica. Section D, Biological crystallography 63, 58-61.

Evans, P.R., 2011. An introduction to data reduction: space-group determination, scaling and intensity statistics. Acta crystallographica. Section D, Biological crystallography 67, 282-292.

Gruba, N., Martinez, J.I.R., Grzywa, R., Wysocka, M., Skorenski, M., Dabrowska, A., Lecka, M., Suder, P., Sienczyk, M., Pyrc, K., Lesner, A., 2018. One Step Beyond: Design of Substrates Spanning Primed Positions of Zika Virus NS2B-NS3 Protease. ACS Med Chem Lett 9, 1025-1029.

Hajduk, P.J., Huth, J.R., Fesik, S.W., 2005. Druggability indices for protein targets derived from NMR-based screening data. Journal of medicinal chemistry 48, 2518-2525.

Harner, M.J., Frank, A.O., Fesik, S.W., 2013. Fragment-based drug discovery using NMR spectroscopy. Journal of biomolecular NMR 56, 65-75.

Headd, J.J., Echols, N., Afonine, P.V., Grosse-Kunstleve, R.W., Chen, V.B., Moriarty, N.W., Richardson, D.C., Richardson, J.S., Adams, P.D., 2012. Use of knowledge-based restraints in phenix.refine to improve macromolecular refinement at low resolution. Acta crystallographica. Section D, Biological crystallography 68, 381-390.

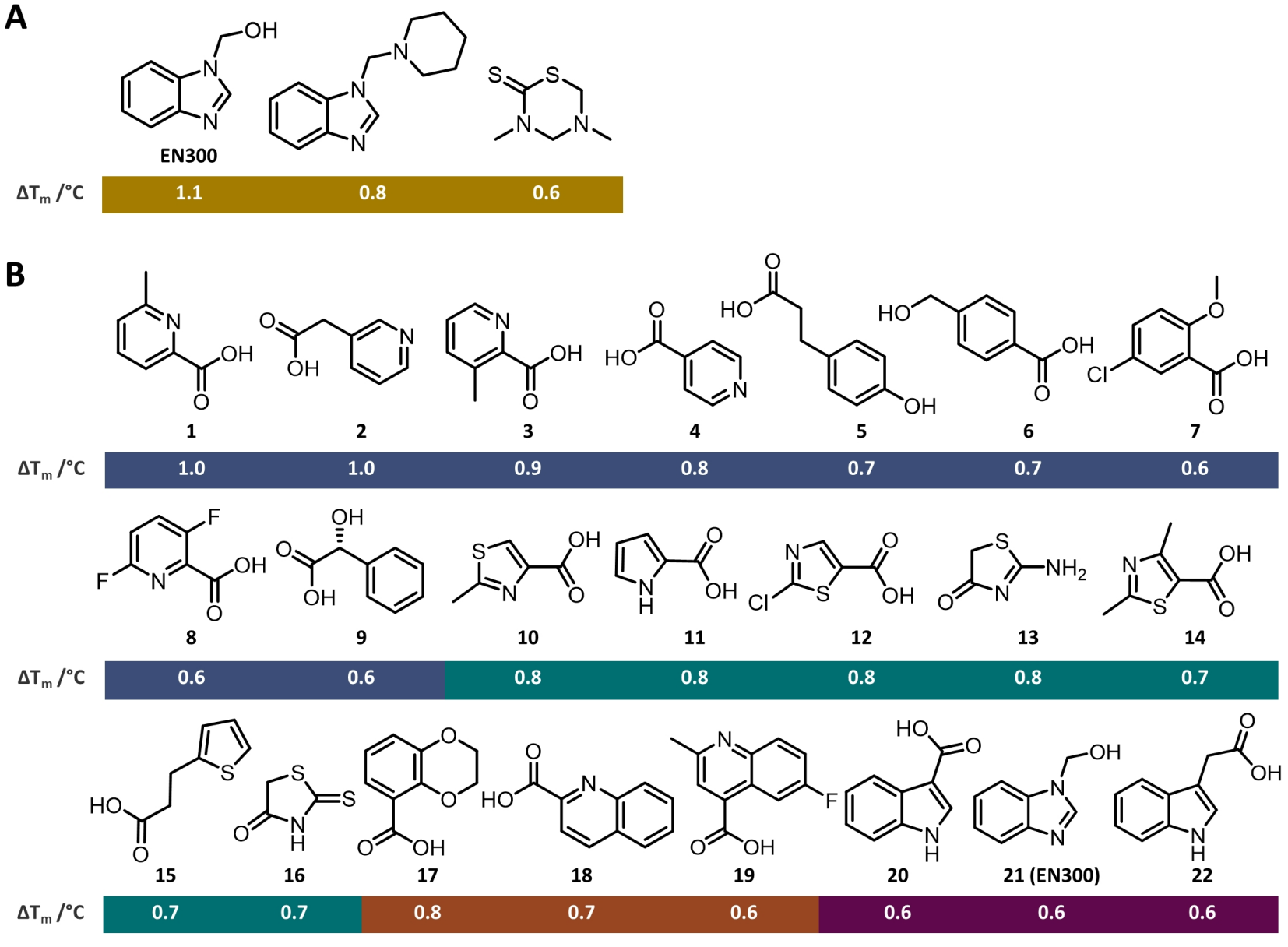
Johnson, B.A., 2004. Using NMRView to visualize and analyze the NMR spectra of macromolecules. Methods Mol Biol 278, 313-352.

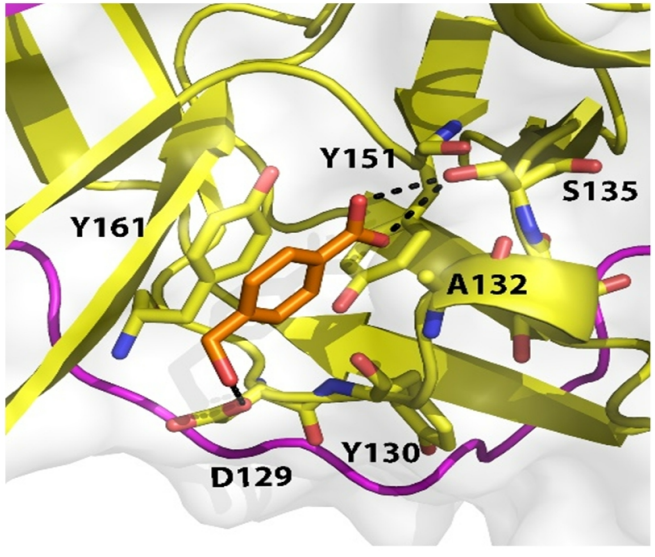
Kang, C., Keller, T.H., Luo, D., 2017. Zika Virus Protease: An Antiviral Drug Target. Trends in microbiology 25, 797-808.

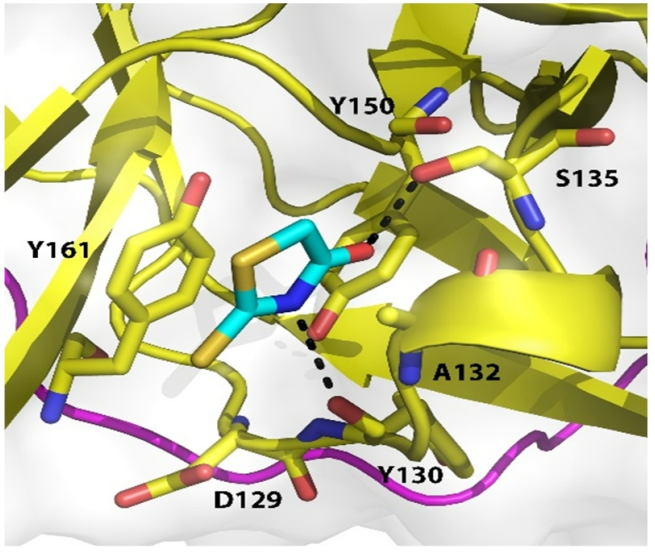
Kim, Y.M., Gayen, S., Kang, C., Joy, J., Huang, Q., Chen, A.S., Wee, J.L., Ang, M.J., Lim, H.A., Hung, A.W., Li, R., Noble, C.G., Lee le, T., Yip, A., Wang, Q.Y., Chia, C.S., Hill, J., Shi, P.Y., Keller, T.H., 2013.

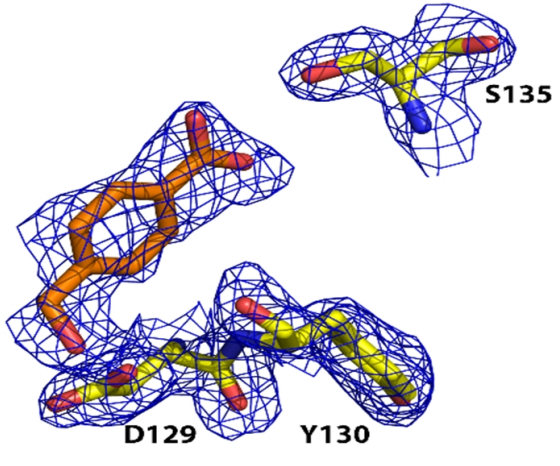
NMR analysis of a novel enzymatically active unlinked dengue NS2B-NS3 protease complex. J Biol Chem 288, 12891-12900.
Koh-Stenta, X., Joy, J., Wang, S.F., Kwek, P.Z., Wee, J.L., Wan, K.F., Gayen, S., Chen, A.S., Kang, C., Lee, M.A., Poulsen, A., Vasudevan, S.G., Hill, J., Nacro, K., 2015. Identification of covalent active site inhibitors of dengue virus protease. Drug Des Devel Ther 9, 6389-6399.
Kumar, A., Liang, B., Aarthy, M., Singh, S.K., Garg, N., Mysorekar, I.U., Giri, R., 2018. Hydroxychloroquine Inhibits Zika Virus NS2B-NS3 Protease. ACS Omega 3, 18132-18141.
Kuno, G., Chang, J.G.-J., 2007. Full-length sequencing and genomic characterization of Bagaza, Kedougou, and Zika viruses. Archives of Virology 152, 687-696.
Lei, J., Hansen, G., Nitsche, C., Klein, C.D., Zhang, L., Hilgenfeld, R., 2016. Crystal structure of Zika virus NS2B-NS3 protease in complex with a boronate inhibitor. Science 353, 503-505.
Li, Y., Loh, Y.R., Hung, A.W., Kang, C., 2018. Characterization of molecular interactions between Zika virus protease and peptides derived from the C-terminus of NS2B. Biochem Biophys Res Commun 503, 691-696.
Li, Y., Phoo, W.W., Loh, Y.R., Zhang, Z., Ng, E.Y., Wang, W., Keller, T.H., Luo, D., Kang, C., 2017. Structural characterization of the linked NS2B-NS3 protease of Zika virus. FEBS Lett 591, 2338-2347.
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.
Luo, D., Vasudevan, S.G., Lescar, J., 2015. The flavivirus NS2B-NS3 protease-helicase as a target for antiviral drug development. Antiviral Res 118, 148-158.
Mahawaththa, M.C., Pearce, B.J., Szabo, M., Graham, B., Klein, C.D., Nitsche, C., Otting, G., 2017. Solution conformations of a linked construct of the Zika virus NS2B-NS3 protease. Antiviral Res 142, 141-147.
Manzano, M., Padia, J., Padmanabhan, R., 2014. Small molecule inhibitor discovery for dengue virus protease using high-throughput screening. Methods Mol Biol 1138, 331-344.
Mass, N., Selisko, B., Malet, H., Peyrane, F., Debarnot, C., Decroly, E., Benarroch, D., Egloff, M.P., Guillernot, J.C., Alvarez, K., Canard, B., 2007. Dengue virus: viral targets and antiviral drugs. Virologie (Montrouge) 11, 121-133.
Mayer, M., Meyer, B., 1999. Characterization of Ligand Binding by Saturation Transfer Difference NMR Spectroscopy. Angewandte Chemie International Edition 38, 1784-1788.
Ndeffo-Mbah, M.L., Parpia, A.S., Galvani, A.P., 2016. Mitigating Prenatal Zika Virus Infection in the Americas. Annals of internal medicine 165, 551-559.
Nitsche, C., Holloway, S., Schirmeister, T., Klein, C.D., 2014. Biochemistry and Medicinal Chemistry of the Dengue Virus Protease. Chemical Reviews.
Nitsche, C., Passioura, T., Varava, P., Mahawaththa, M.C., Leuthold, M.M., Klein, C.D., Suga, H., Otting, G., 2019. De Novo Discovery of Nonstandard Macrocyclic Peptides as Noncompetitive Inhibitors of the Zika Virus NS2B-NS3 Protease. ACS Med Chem Lett 10, 168-174.
Noble, C.G., Shi, P.Y., 2012. Structural biology of dengue virus enzymes: Towards rational design of therapeutics. Antiviral Res 96, 115-126.
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., 2016. Structure of the NS2B-NS3 protease from Zika virus after self-cleavage. Nat Commun 7, 13410.
Phoo, W.W., Zhang, Z., Wirawan, M., Chew, E.J.C., Chew, A.B.L., Kouretova, J., Steinmetzer, T., Luo, D., 2018. Structures of Zika virus NS2B-NS3 protease in complex with peptidomimetic inhibitors. Antiviral Research 160, 17-24.
Potterton, E., Briggs, P., Turkenburg, M., Dodson, E., 2003. A graphical user interface to the CCP4 program suite. Acta crystallographica. Section D, Biological crystallography 59, 1131-1137.
Poulsen, A., Kang, C., Keller, T.H., 2014. Drug design for flavivirus proteases: what are we missing? Curr Pharm Des 20, 3422-3427.
Roy, A., Lim, L., Srivastava, S., Lu, Y., Song, J., 2017. Solution conformations of Zika NS2B-NS3pro and its inhibition by natural products from edible plants. PloS one 12, e0180632.

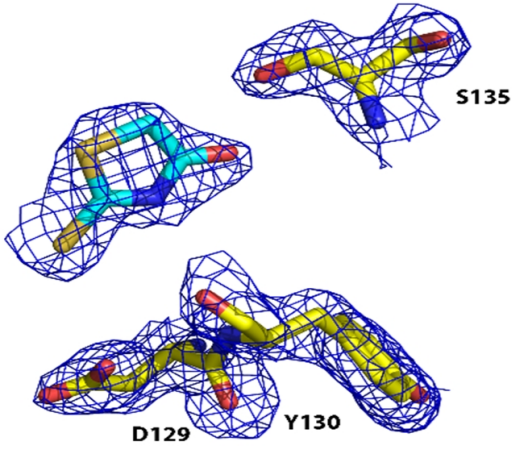
Schöne, T., Grimm, L.L., Sakai, N., Zhang, L., Hilgenfeld, R., Peters, T., 2017. STD-NMR experiments identify a structural motif with novel second-site activity against West Nile virus NS2B-NS3 protease. Antiviral Research 146, 174-183.
Shiryaev, S.A., Farhy, C., Pinto, A., Huang, C.-T., Simonetti, N., Ngono, A.E., Dewing, A., Shresta, S., Pinkerton, A.B., Cieplak, P., Strongin, A.Y., Terskikh, A.V., 2017. Characterization of the Zika virus two-component NS2B-NS3 protease and structure-assisted identification of allosteric small-molecule antagonists. Antiviral Research 143, 218-229.
Skora, L., Jahnke, W., 2017. 19F-NMR-Based Dual-Site Reporter Assay for the Discovery and Distinction of Catalytic and Allosteric Kinase Inhibitors. ACS Medicinal Chemistry Letters 8, 632-635.
Su, X.C., Ozawa, K., Qi, R., Vasudevan, S.G., Lim, S.P., Otting, G., 2009. NMR analysis of the dynamic exchange of the NS2B cofactor between open and closed conformations of the West Nile virus NS2B-NS3 protease. PLoS Negl Trop Dis 3, e561.
Uchil, P.D., Satchidanandam, V., 2003. Architecture of the Flaviviral Replication Complex: PROTEASE, NUCLEASE, AND DETERGENTS REVEAL ENCASEMENT WITHIN DOUBLE-LAYERED MEMBRANE COMPARTMENTS. Journal of Biological Chemistry 278, 24388-24398.
Viegas, A., Manso, J., Nobrega, F.L., Cabrita, E.J., 2011. Saturation-Transfer Difference (STD) NMR: A Simple and Fast Method for Ligand Screening and Characterization of Protein Binding. Journal of Chemical Education 88, 990-994.
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., 2019. Discovery, X-ray Crystallography and Antiviral Activity of Allosteric Inhibitors of Flavivirus NS2B-NS3 Protease. J Am Chem Soc 141, 6832-6836.
Yildiz, M., Ghosh, S., Bell, J.A., Sherman, W., Hardy, J.A., 2013. Allosteric Inhibition of the NS2B-NS3 Protease from Dengue Virus. ACS Chem Biol 17, 76-80.
Zhang, Z., Li, Y., Loh, Y.R., Phoo, W.W., Hung, A.W., Kang, C., Luo, D., 2016. Crystal structure of unlinked NS2B-NS3 protease from Zika virus. Science 354, 1597-1600.


A
EN300
ΔTm /°C
1.1
0.8
0.6
B
1
2
3
4
5
6
7
ΔTm /°C
1.0
1.0
0.9
0.8
0.7
0.7
0.6
8
9
10
11
12
13
14
ΔTm /°C
0.6
0.6
0.8
0.8
0.8
0.8
0.7
15
16
17
18
19
20
21 (EN300)
22
ΔTm /°C
0.7
0.7
0.8
0.7
0.6
0.6
0.6
0.6



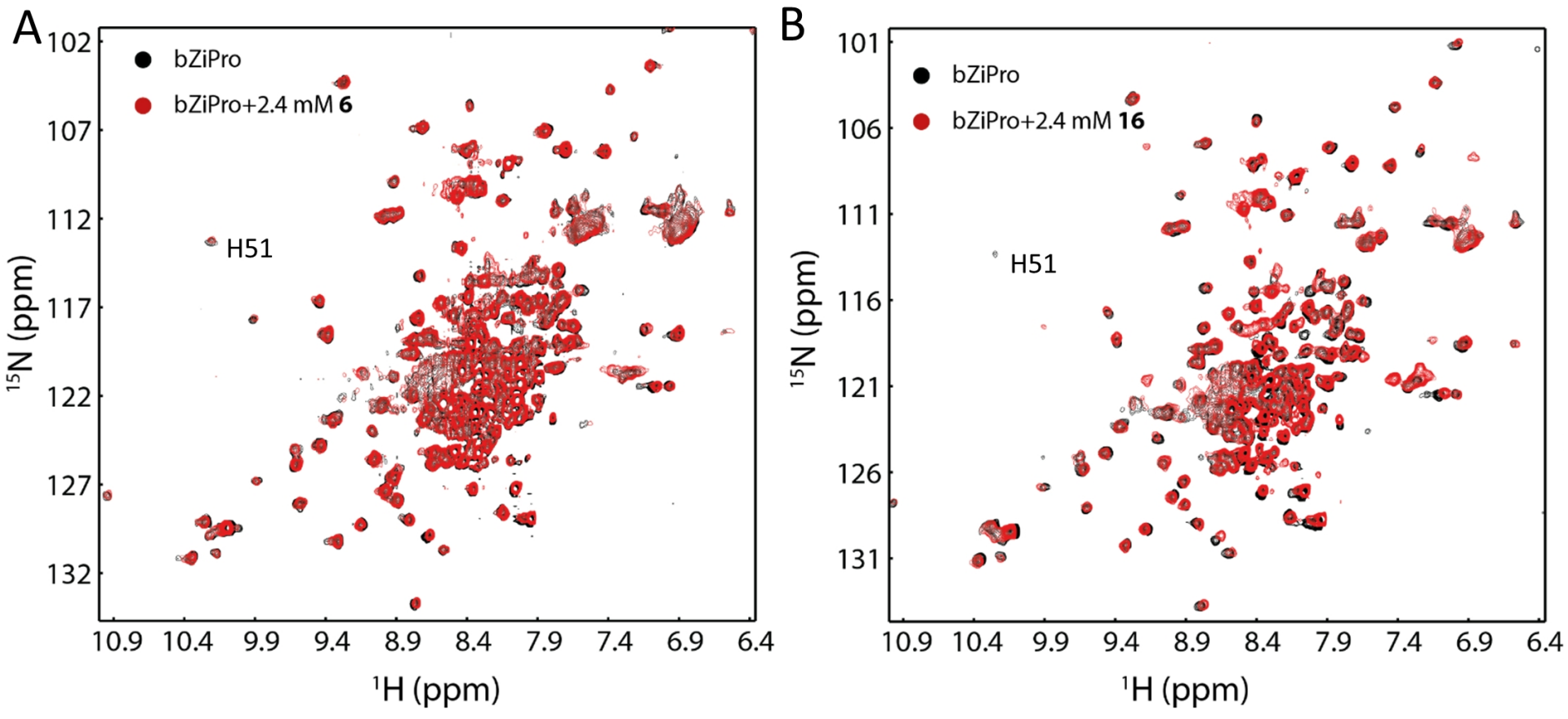


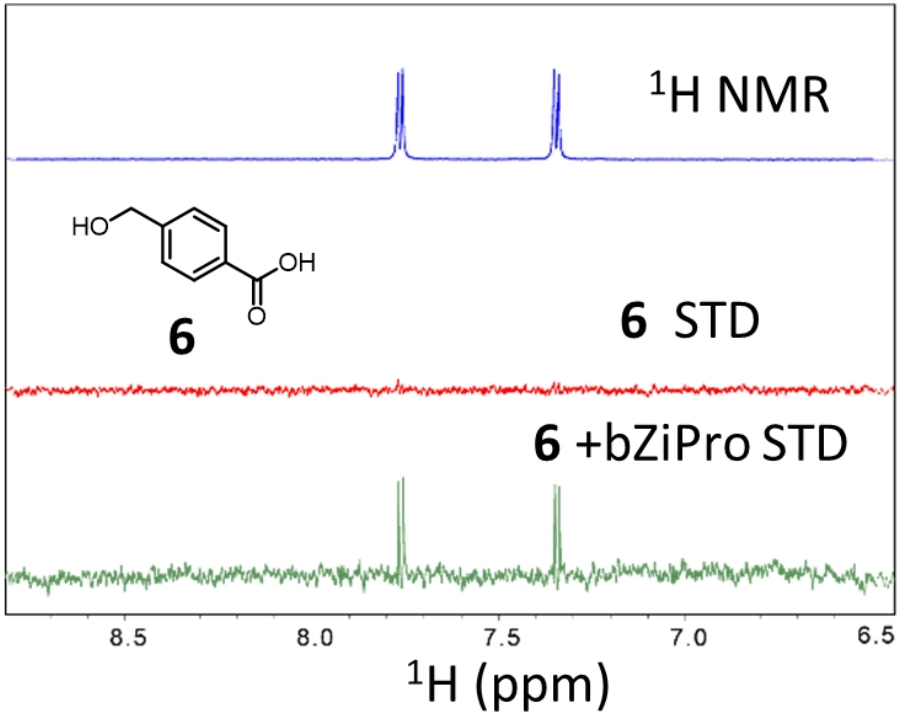




Compound **6**

Compound **16**


A
bZiPro
bZiPro+2.4 mM 6
H51
^{15}N (ppm)
^{1}H (ppm)
B
bZiPro
bZiPro+2.4 mM 16
H51
^{15}N (ppm)
^{1}H (ppm)


^{1}H NMR
HO
OH
O
6
6 STD
6 +bZiPro STD
8.5
8.0
7.5
7.0
6.5
^{1}H (ppm)

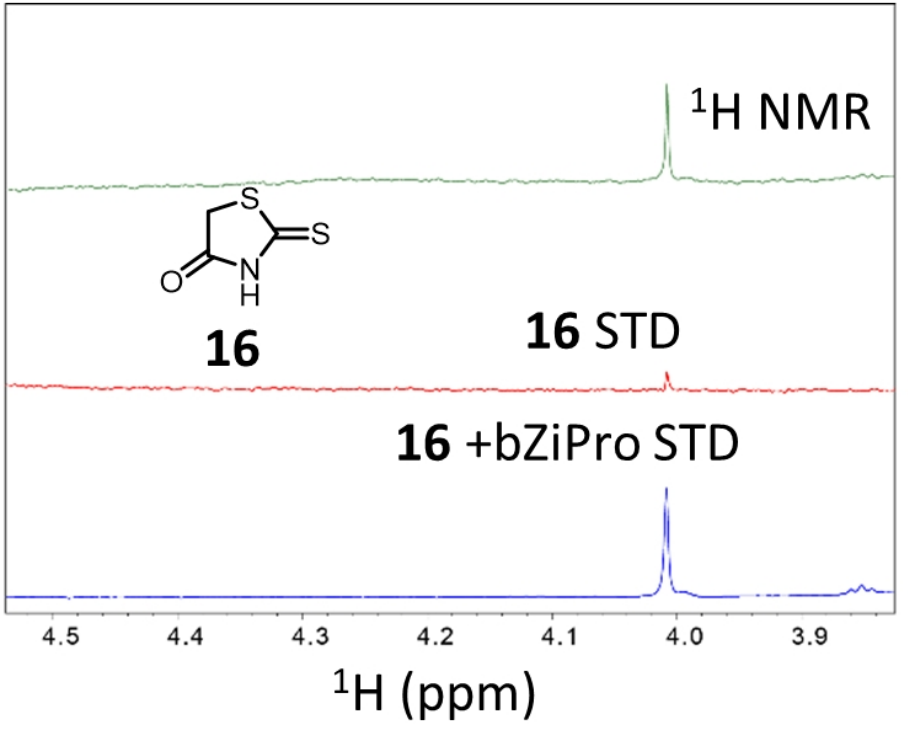

^{1}H NMR
S
S
O
N
H
16
16 STD
16 +bZiPro STD
4.5
4.4
4.3
4.2
4.1
4.0
3.9
^{1}H (ppm)

Table S1. X-ray data collection and refinement statistics

Table 1. Data Collection and Refinement Statistics		
Data Collection Statistics	bZiPro + Compound 6 (PDB Code: 6L4Z)	bZiPro + Compound 16 (PDB Code: 6L50)
Wavelength (Å)	0.9537	0.9537
Resolution range (Å)	45.76 - 1.9 (1.968 - 1.9)	42.3 - 1.95 (2.02 - 1.95)
Space group	P 21 21 21	P 21 21 21
Unit cell a, b, c	59.549 59.525 214.556	59.744 59.909 214.833
α, β, γ (Å) (°)	90 90 90	90 90 90
Total reflections	370115 (36242)	418435 (41495)
Unique reflections	61183 (6015)	56428 (5506)
Multiplicity	6.0 (6.0)	7.4 (7.5)
Completeness (%)	99.93 (100.00)	98.46 (97.62)
Mean I/sigma (I)	14.09 (1.69)	27.73 (3.20)
[a]R_{merge}	0.5746 (1.109)	0.2506 (0.647)
R-meas	0.6247 (1.215)	0.2698 (0.6946)
R-pim	0.2429 (0.4897)	0.09903 (0.2515)
CC1/2	0.812 (0.27)	0.977 (0.586)
CC*	0.947 (0.652)	0.994 (0.86)
Refinement Statistics		
Reflections used in refinement	61152 (6015)	56428 (5505)
Reflections used for R-free	3012 (293)	1999 (194)
[b]R_{work}	0.1785 (0.2701)	0.2149 (0.3763)
[c]R_{free}	0.2244 (0.2918)	0.2348 (0.3575)
Number of non-hydrogen atoms	5807	5707
macromolecules	5622	5629
ligands	11	7
solvent	174	71
Protein residues	761	760
[d]RMSD (bonds) (Å)	0.003	0.004
RMSD (angles) (°)	0.58	0.67
Ramachandran favored (%)	95.55	96.75
Ramachandran allowed (%)	4.32	3.25
Ramachandran outliers (%)	0.13	0.00
Rotamer outliers (%)	0.18	0.00
Clashscore	6.32	7.86
Average B-factor	32.29	31.63
macromolecules	32.32	31.70

ligands	34.67	35.19
solvent	31.31	26.25

Statistics for the highest-resolution shell are shown in parentheses.

[a]$R_{merge} = \sum|I_j - < I > |/\sum I_j$, where I_j is the intensity of an individual reflection, and $< I >$ is the average intensity of that reflection.

[b]$R_{work} = \sum||F_{obs}| - |F_{calc}||/\sum|F_{obs}|$, where F_{obs} 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% of randomly chosen reflections omitted from the refinement.

[d]RMSD, root mean square deviations.

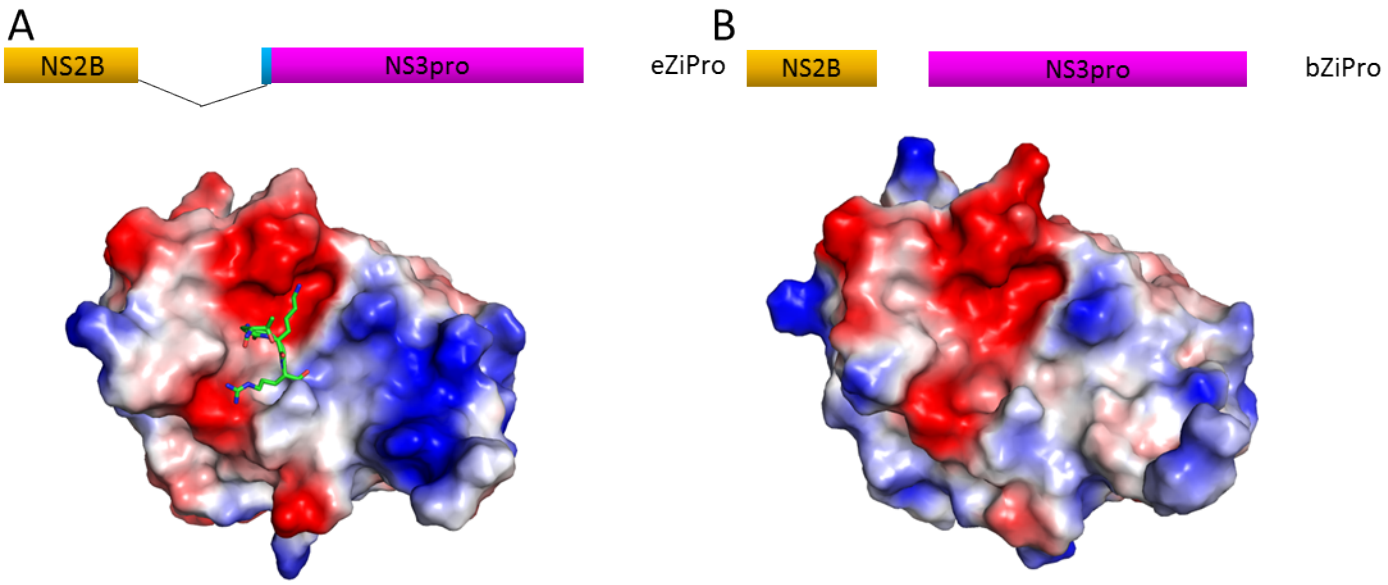


Figure S1 Protease constructs used in this study. A. eZiPro protease construct. This construct contains the native linker of NS2B and NS3 while the transmembrane domains of NS2B are not included. Crystal structure of eZiPro (PDB id 5GJ4) [1] is shown the peptide in the active pocket is shown in sticks. B. Crystal structure of bZiPro. The structure of bZiPro (PDB id 5GPI) is shown and the pocket is empty [2].

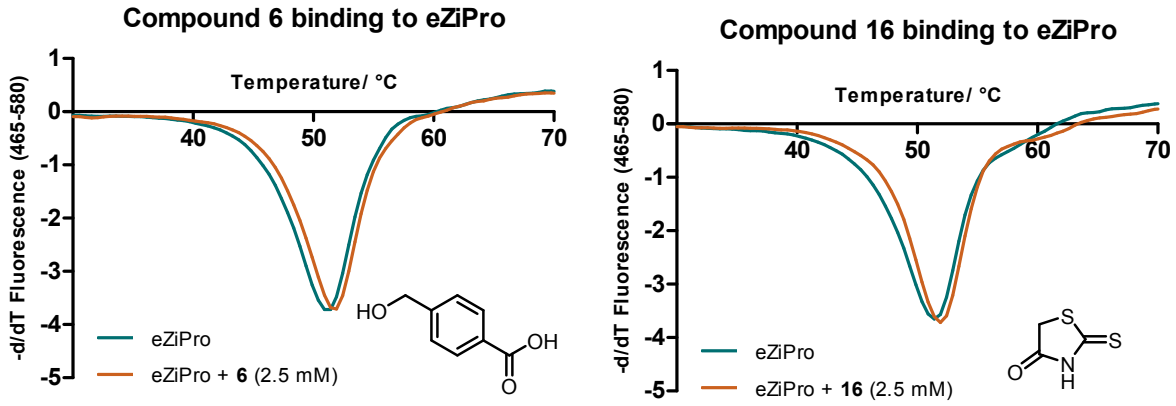
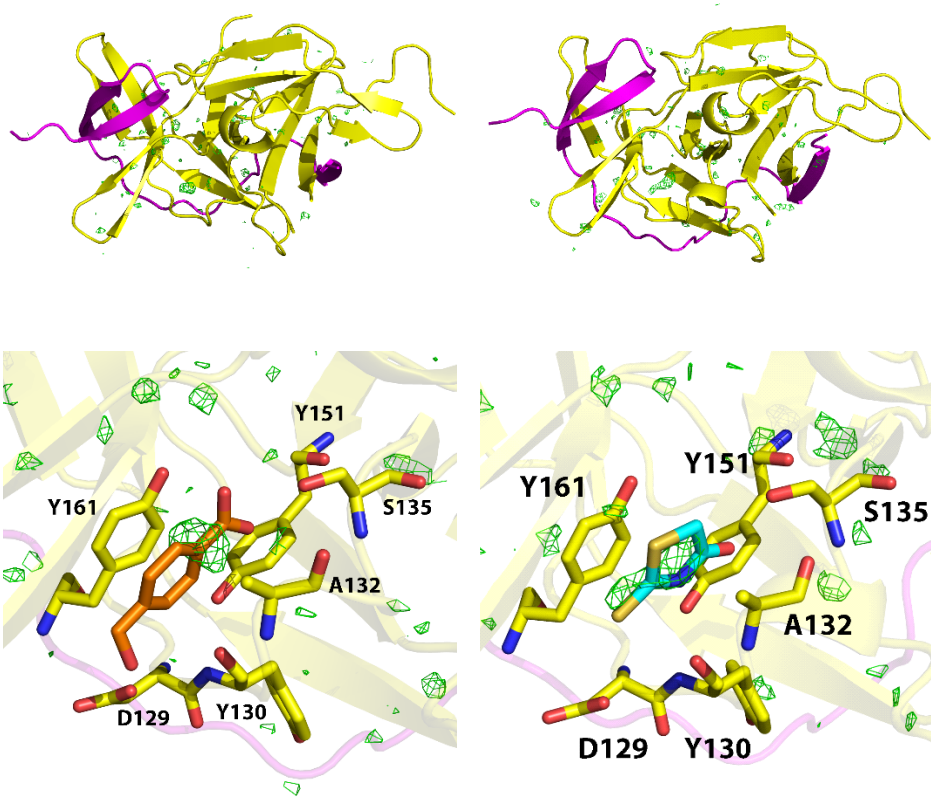


Figure S2 Thermal shift assay of compounds **6** and **16**.



Compound **6** Compound **16**

Figure S3 Omit map of bZiPro in complex with compound 6 and 16. The upper panel is the omit maps (mFo-Fc) coloured in green and contoured at 3σ. The lower panel is the closed-up view of the omit maps at the active site. The neighboring residues are shown as sticks in yellow.

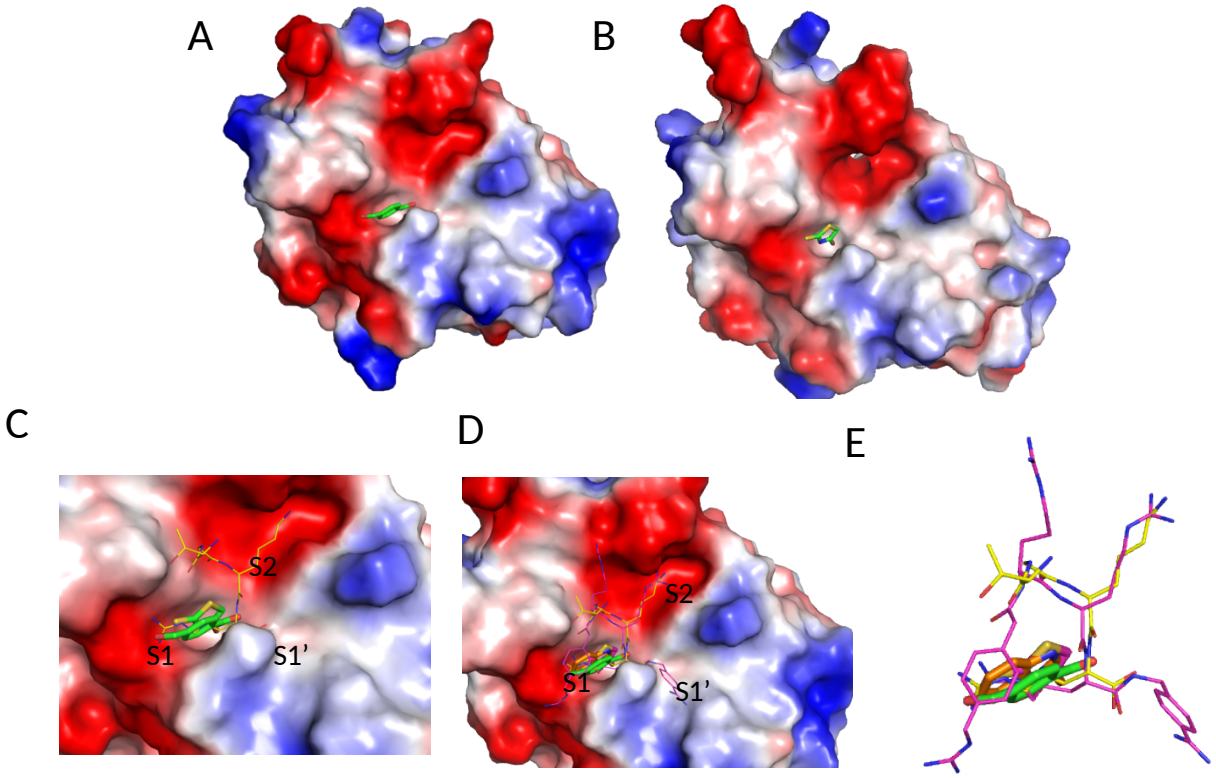


Figure S4 Comparison of the structures of ZIKV protease in complex with different types of ligands/inhibitors A,B. Structures of bZiPro in complexes with fragment compounds **6** and **16**. C. Overlay of the fragments with TGKR peptide in eZiPro. The structure of eZiPro [1] (PDB id 5GJ4) is shown. The TGKR peptide is shown in line and the fragments are shown as sticks, respectively. D. Overlay of ligands/inhibitors forming complex with ZIKV protease. E. Overlay of the inhibitors. Structures of ZIKA in complex with TGKR (PDB id 5GJ4) and a peptide inhibitor (PDB id 5ZMQ) are shown.

References

[1] 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, Nat Commun, 7 (2016) 13410.
[2] Z. Zhang, Y. Li, Y.R. Loh, W.W. Phoo, A.W. Hung, C. Kang, D. Luo, Crystal structure of unlinked NS2B-NS3 protease from Zika virus, Science, 354 (2016) 1597-1600.