Sussex Research Online

Structural insights into the inhibition of Zika virus NS2B-NS3 protease by a small-molecule inhibitor

Article (Accepted Version)

Li, Yan, Zhang, Zhenzhen, Phoo, Wint Wint, Loh, Ying Ru, Li, Rong, Yang, Hai Yan, Jansson, Anna E, Hill, Jeffrey, Keller, Thomas H, Nacro, Kassoum, Luo, Dahai and Kang, CongBao (2018) Structural insights into the inhibition of Zika virus NS2B-NS3 protease by a small-molecule inhibitor. Structure, 26 (4). pp. 555-564. ISSN 0969-2126

This version is available from Sussex Research Online: http://sro.sussex.ac.uk/id/eprint/87884/

This document is made available in accordance with publisher policies and may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the URL above for details on accessing the published version.

Copyright and reuse:
Sussex Research Online is a digital repository of the research output of the University.

Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable, the material made available in SRO has been checked for eligibility before being made available.

Copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

http://sro.sussex.ac.uk

Structural insights into the inhibition of Zika virus NS2B-NS3 protease by a small molecule inhibitor

Yan Li[1,†], Zhenzhen Zhang[2,3,†], Wint Wint Phoo[2,3,4,†], Ying Ru Loh[1], Rong Li[1], Hai Yan Yang[1], Anna E Jansson[1], Jeffrey Hill[1], Thomas H. Keller[1], Kassoum Nacro[1*], Dahai Luo[2,3*], CongBao Kang[1*]

Affiliations

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

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

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

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

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

† These authors contributed equally to the work;

Summary

Zika virus (ZIKV) infection has become a global public health concern. The viral NS2B-NS3 protease is an attractive antiviral target because of its role in maturation of viral non-structural proteins. Substrate-derived protease inhibitors have been investigated while it remains challenging to develop them into drugs. Small molecule inhibitors are of great interest in antiviral drug development. Here we report the structure and dynamics of ZIKV NS2B-NS3 protease covalently bound to a small molecule inhibitor. Our crystallographic and NMR studies have demonstrated that the inhibitor further stabilizes the closed conformation of ZIKV protease. Upon hydrolyzed *in situ* into two fragments, the benzoyl group of the inhibitor forms a covalent bond with the side chain of catalytic residue S135 and the second fragment exhibits no obvious molecular interactions with the protease. This study provides a detailed mechanism of action for the covalent inhibitor, which will guide further development of ZIKV protease inhibitors.

Keywords: Zika virus; protease; drug discovery; protein dynamics; structure; protease inhibitor

Introduction

Zika virus (ZIKV) is a mosquito-borne pathogen and a member of the *Flaviviridae* family which also includes other important human pathogens such as Dengue virus (DENV), West Nile virus (WNV), yellow fever virus, Murray Valley encephalitis virus, and Japanese encephalitis virus. The recent outbreak of ZIKV has threatened global populations. ZIKV infection is found to be related to neurological disorders such as Guillain–Barré syndrome, acute myelitis (Broutet et al., 2016; Petersen et al., 2016) and microcephaly (Ndeffo-Mbah et al., 2016). Therefore, effective strategies such as vaccines and antivirals are urgently needed. Effort has been made to develop chemical therapies against its close pathogens such as DENV and WNV (Jones et al., 2015; Luo et al., 2015), but there are still no clinically approved antivirals available.

The ZIKV genome encodes a polyprotein that can be post-translationally processed by both host signal peptidases and viral protease into three structural proteins (C, prM, and E) required for viral particle assembly and seven non-structural proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5) required for viral replication and virion assembly. The viral protease is a two-component serine protease formed by the N-terminal part of NS3 and the cofactor region of NS2B-a membrane protein essential for the membrane location of NS3 (Baronti et al., 2014; Clum et al., 1997; Li et al., 2015). This two-component complex forms an active site with a catalytic triad composed of H51, D75, and S135 (Phoo et al., 2016; Zhang et al., 2016). The viral protease cleaves the junctions of NS2A/NS2B, NS2B/NS3, NS3/NS4A, and NS4B/NS5 to release functional proteins. Structural studies have demonstrated that the protease recognizes the sites composed of positive charged residues such as R and K at the P1 and P2 positions, respectively (Lei et al., 2016; Phoo et al., 2016). The negative charges at the protease active site make it challenging to develop small molecule substrate competitive inhibitors (Kang et al., 2017; Poulsen et al., 2014).

Different approaches such as high-throughput screening (HTS), structure-based drug design, substrate-based inhibition, and *in silico* docking have been used to develop DENV and WNV protease inhibitors (Deng et al., 2012; Lim and Shi, 2013; Lim et al., 2013; Nitsche et al., 2012; Yang et al., 2011; Yin et al., 2006). Although no DENV or WNV protease inhibitor has entered clinical stages, some of the available inhibitors have been shown to be active against ZIKV protease (Lei et al., 2016; Li et al., 2017b). This is not surprising because DENV, WNV, and ZIKV proteases share high structure similarities (Kang et al., 2017). The structure of ZIKV NS2B-NS3 protease is almost identical to those of DENV and WNV proteases (Chen et al., 2016; Erbel et al., 2006; Lei et al., 2016; Phoo et al., 2016; Zhang et al., 2016). We have used an unlinked ZIKV protease construct (designated as bZiPro) for structural and inhibitor binding studies (Li et al., 2017a; Li et al., 2017b; Phoo et al., 2016; Zhang et al., 2016). Although conformational exchanges exist in bZiPro, it is dominantly in the closed conformation and is more suitable for structural and ligand binding studies than the conventional linked construct (Kang et al., 2017; Li et al., 2017a). Structural information for ZIKV protease in complex with several inhibitors is available, which supports antiviral inhibitor design (Lei et al., 2016; Phoo et al., 2016; Zhang et al., 2016). Peptidic inhibitors such as Ac-Lys-Arg-H derived from protease substrates are potent competitive inhibitors, but it is very challenging to develop them into drug-like molecules due to issues such as cell penetration and stability (Li et al., 2017b), which has stimulated the search for new inhibitors (Luo et al., 2015; Nitsche et al., 2014b). Small molecule inhibitors are of great interest. To date, there is only one structure available for ZIKV protease in complex with a non-pepditic chemical fragment is available (Zhang et al., 2016). Lack of ZIKV protease structures in complex with small molecular weight inhibitors limits the understanding of the molecular interaction and structure-based drug design (Zhang et al., 2016). Pyrazole ester derivatives are active against both WNV and DENV proteases and identified from a HTS of a small molecule library from NIH (Johnston et al., 2007). This type of inhibitors are unstable in solution (Sidique et al., 2009) and have been shown to covalently

link to DENV protease (Koh-Stenta et al., 2015). However, the structural information of the complex is not well described due to the difficulty in structural studies (Koh-Stenta et al., 2015).

In this study, we demonstrate that bZiPro forms a covalent complex with one pyrazole ester derivative, 5-amino-1-((4-methoxyphenyl)sulfonyl)-1H-pyrazol-3-yl benzoate (compound **1**) evidenced by the mass spectrometry. We then solved the crystal structure of bZiPro in complex with compound **1**, which reveals that the benzoyl moiety of the inhibitor forms a covalent bond with S135 of NS3. The remaining hydroxyl-pyrazole moiety from compound **1** is released into the media and is not found bound to the protease, which is also verified by NMR binding study. The crosslinked benzoyl group further stabilizes the closed conformation of bZiPro as NMR signals of most residues from both NS2B and NS3 are well resolved. Further binding study demonstrated that the integrity of the compound is critical for protease binding as the fragments derived from compound **1** exhibited no interaction with bZiPro. Compound **5,** a derivative of compound **1,** interacts similarly with and inhibits bZiPro strongly. Our study provides the structural information of ZIKV protease in complex with a covalent small molecule to understand the covalent interactions. Our current study is therefore useful for structure-based design of ZIKV protease inhibitors.

Results

Compound 1 reacts with ZIKV NS2B-NS3 protease

Compound **1** was considered as an allosteric inhibitor of WNV NS2B-NS3 protease (Sidique et al., 2009) and a further study has shown that it forms covalent interactions with DENV protease (Koh-Stenta et al., 2015) (**Fig. 1A**). We found that compound **1** inhibits the enzymatic activity of bZiPro with an IC_{50} of 1.5 μM when Benzoyl-Nle-Lys-Arg-Arg-aminomethylcoumarin (Bz-nKRR-AMC) was used as a substrate. bZiPro exhibited a 104 Da molecular weight increase in the ESI-TOF MS spectrum after treatment with compound

1(**Fig. 1B**), mirroring the results from DENV protease and indicating the modification of bZiPro by the benzoyl moiety based on the structure of compound **1** (Koh-Stenta et al., 2015) (**Fig. 1C**). The ester group in the inhibitor makes compound **1** unstable in solution, giving rise to production of the benzoic acid and 5-amino-1-((4-methoxyphenyl)sulfonyl)-1*H*-pyrazol-3-ol (compound **2**) (Sidique et al., 2009).

Crystal structure of bZiPro (C143S)-compound 1 complex

Zika protease is able to form a dimer and a disulfide bond between C143 residues of neighboring NS3 in the crystal was observed (Kang et al., 2017). To remove the crystallographic disulfide bond of C143, A C143S mutation was generated and the resulting mutant has little influence on the protease activity (Figure S1). We determined the crystal structure of bZiPro (C143S)-compound **1** complex at a resolution of 1.9 Å (**Fig. 2, Table 1**). The four protease molecules in one unit cell were labeled following the peptide chain IDs (Figure S1). Consistent with the MS result, the benzoyl moiety of compound **1** forms a covalent bond with the side chain of S135. The benzoyl moiety sits in the S1 pocket and overlays with the aliphatic chain of the P1 residue of the Ac-Lys-Arg-H inhibitor (**Fig. 2**). Atoms of several residues are in close contact (less than 5 Å) with the benzoyl moiety. There is one hydrogen bond formed between the oxygen atom and the nitrogen atom of the side chain of H51, which may be critical for the closed conformation of the complex (**Fig. 2**). The benzoyl moiety of the compound **1** also forms π- π stacking interactions between Y161 and A132. The compound **1** does not have molecular interactions with the backbone amide protons of G133 and T134. The oxyanion hole is also not formed due to the cis conformation of the A132-G133 peptide bond (**Fig. 2B**). The remaining portion (compound **2**) of compound **1** is not present in the structure, suggesting its release from the active site upon trans-esterification reaction by S135. Unlike the NS2B-NS3 cleavage site peptide, Thr-Gly-Lys-Arg (TGKR) and other inhibitors, benzoyl moiety does not form any interaction with residues of NS2B, suggesting that the occupancy of the S1 pocket alone is sufficient for suppressing the open conformation of protease (Phoo et al., 2016).

Crystal structure of bZiPro (C143S)-Ac-Lys-Arg-H complex

We also determined the crystal structure of bZiPro (C143S)-Ac-Lys-Arg-H complex to a high resolution of 1.51 Å with Rwork and Rfree values of 18.78 % and 20.16 % respectively (**Fig. 3**, Table 1). The protease in the complex adopts a very similar structure to the WT reported recently (**Fig. 3**) (Li et al., 2017b). At 1.51 Å resolution, the electron density of the aldehyde moiety of the inhibitor and the surrounding protease residues of bZiPro (C143S) are improved significantly compared to the WT structure. We can observe that the peptide bond between A132 and G133 adopt two conformations and the oxyanion hole is also not formed in bZiPro (C143S) structure (**Fig. 3C, D**). The conformation of the hemiacetal formed between Ac-Lys-Arg-H and S135 can also be refined clearly, allowing the comparison with the ester bond formed between the benzoyl group and S135 (**Fig. 3**).

Unique adduct of the benzoyl moiety to S135

The structure of NS3 and the N-terminal part of NS2B in the complex adopts a similar conformation to those of the linked and unlinked protease in the free form and complexes with substrate and inhibitors. In the complex, the protease is in the closed conformation in which the C-terminal region (residues 71-87) of NS2B engages with NS3 and forms part of the active site (**Fig. 2**). The root-mean-squared deviations (rmsd) of the Cα atoms of observed residues are 0.36 Å (PDB 5GPI, (Zhang et al., 2016)), 0.26 Å (PDB id 5H4I, (Zhang et al., 2016)), 0.41 Å (PDB id 5GJ4, (Phoo et al., 2016)), 0.29 Å (PDB id 5H6V, (Li et al., 2017b)), and 0.61 Å (PDB id 5LC0, (Lei et al., 2016)) compared with those of bZiPro in the absence (Zhang et al., 2016) and presence of EN300 compound (Li et al., 2017b), TGKR peptide (Phoo et al., 2016), Ac-Lys-Arg-H dipeptide (Li et al., 2017b), and a boronic peptide inhibitor (Lei et al., 2016), respectively (Figure S2, Table S1). To further probe the molecular interactions, the ^{1}H-^{15}N-HSQC spectra of bZiPro in the absence and presence of compound **1** were collected and compared. Compared with free protein, bZiPro in complex with compound **1** exhibits profound chemical shift changes in the ^{1}H-^{15}N-HSQC spectrum, which clearly

indicates their molecular interactions in solution (**Fig. 4A**). In the free ZIKV protease, conformational exchanges between the C-terminal β-hairpin region and NS3 broadened the cross peaks of corresponding residues in the ^{1}H-^{15}N-HSQC spectrum (Zhang et al., 2016). Most of the residues exhibit detectable cross peaks in the ^{1}H-^{15}N-HSQC spectrum of the complex and can be unambiguously assigned (Figure S3), suggesting that compound **1** binding stabilizes the protease conformation in solution without direct interactions with residues from NS2B. In addition, compared with the available assignments of free bZiPro, several residues exhibit profound chemical shift perturbations upon inhibitor binding and are limited to the vicinity of the catalytic triad, suggesting that compound **1** binds to the active site (**Fig. 3B**). Although chemical shift difference was observed for the ^{1}H-^{15}N spectra of bZiPro in complexes with compound **1** and Ac-Lys-Arg-H, the numbers of the cross peaks of corresponding spectra are similar, indicating that bZiPro is in the closed conformation in both complexes. The observed chemical shift difference is due to the different binding modes of these two inhibitors (**Fig. 2**, Figure S4)

Structure and dynamics of bZiPro-compound 1 complex in solution

Compound **1** was titrated into bZiPro and the result showed that several residues such as L98 and L140 were undergoing slow exchanges in NMR time scale (**Fig. 4C, D, E**), indicative of their monovalent and strong bindings in solution. The result agrees well with the conclusion derived from MS analysis and X-ray structure in which compound **1** has covalent interactions with bZiPro, which was also observed for both West Nile and Dengue virus proteases (Koh-Stenta et al., 2015).

^{15}N-spin-lattice relaxation rate (R_1), spin-spin relaxation rate (R_2), and steady-state heteronuclear NOEs (hetNOE) (Kay et al., 1989) values of the complex were obtained to understand its dynamics in solution (**Fig. 5A**). The average ratios of R_2/R_1 values of N-terminal region of NS2B cofactor region, the β-hairpin region of NS2B, and NS3 excluding flexible residues are very similar (**Fig. 5A**). The N-terminal region of the

NS2B cofactor is engaged with NS3 regardless of ligand binding while the β-hairpin region of NS2B has different conformations in the absence of an inhibitor. The R_1 and R_2 values of residues from structured regions of NS2B and NS3 in the complex are almost identical, corroborating the conclusion that bZiPro is in the closed conformation and NS2B cofactor forms a stable complex with NS3 (**Fig. 5A**). The termini of NS2B and NS3 are highly dynamic in solution because of the low hetNOEs values. The high mobility of the terminus of NS3 in the absence and presence of compound **1** indicates that the corresponding residues do not form stable interactions with the S prime site of the protease. Other residues in bZiPro of the complex exhibit high hetNOEs (>0.6) values, indicative of their rigid structures in solution. Overall, bZiPro-compound **1** forms a stable structure in solution.

The following evidence derived from NMR data support the conclusion obtained from the X-ray structure that the protease in the complex adopts the closed conformation. Firstly, nearly complete resonances from bZiPro in the ^{1}H-^{15}N-HSQC spectrum have been assigned, suggesting the uniform structure of the bZiPro in the complex as signal broadening of residues in the free protease was suppressed by inhibitor binding (Figure S3). Similar result was obtained for WNV protease where the closed conformation made it possible to assign most of the resonances in the ^{1}H-^{15}N-HSQC spectrum (Su et al., 2009). Secondly, secondary structure analysis using TALOS based on the backbone chemical shifts demonstrates the presence of β-sheet structures in the C-terminal region of the NS2B cofactor region (Figure S5). In the open form, the region is unstructured and away from the active site (Jones et al., 2015). The presence of β-sheet structures in solution suggests the formation of the closed conformation in the complex (Figure S5). Thirdly, relaxation analysis shows that most residues in bZiPro form rigid structure in solution (**Fig. 5A**) and residues from NS2B β-hairpin region exhibit similar dynamic values to those of NS3, which confirms that these two regions form a stable complex in solution. Lastly, unambiguous intermolecular NOEs between amide proton of residues K117 of NS3 and amide protons of V76 and L74 of NS2B are identified, supporting the conclusion that NS2B cofactors forms a complex with NS3 and the complex is in the closed

conformation (**Fig. 5B**). In addition, NOEs between amide protons of residues D83 and D79 are identified, which further confirms the presence of the β-hairpin structure in the bZiPro-compound **1** complex (**Fig. 5B, C**).

The integrity of compound 1 is essential for protease binding

MS result and X-ray structure indicate that compound **1** reacts with bZiPro in solution as only the benzoyl moiety is covalently-linked to bZiPro. The products of the trans-esterification reaction between bZiPro and compound **1** are the benzoyl moiety linked to S135 of bZiPro and compound **2** (**Fig. 1**). Ultimately, the complex benzoyl-bZiPro might be the subject of a hydrolysis to give a benzoic acid derivative and regenerate the protease S135 hydroxyl while the benzoyl modified bZiPro was stable in solution during our NMR data collection. We then investigated whether benzoic acid and compound **2** can bind to bZiPro, respectively. ^{1}H-^{15}N-HSQC spectra of bZiPro in the absence and presence of these two compounds were collected and compared (Figure S6). Interestingly, both compounds do not exhibit obvious interactions with bZiPro as no profound chemical shift differences were observed in the spectra (Figure S6), suggesting that the integrity of compound **1** is required for the activity. Compound **4** has an ester group and a benzoyl moiety and exhibits no obvious interactions with bZiPro (Figure S6), suggesting that compound **1** binding to the active site is prerequisite for the covalent modification and the pyrazole moiety plays an important role in compound **1** binding to bZiPro prior to S135 hydroxyl attack.

Stronger inhibition with a biphenyl moiety from compound 5

To test the idea that modified benzoyl moiety may alter the strength of inhibition, compound **5** with a biphenyl ester rather than phenyl ester group (compound **1**) was synthesized. The IC_{50} of compound **5** on bZiPro is 0.1 μM (**Fig 6A**), exhibiting an improvement over compound **1** whose IC_{50} on bZiPro was 1.5 μM. Molecular weight of bZiPro in complex with compound **5** was subjected to MS analysis. As expected, NS3 exhibits a molecular weight increase of 214.6 Da in the spectrum (**Fig. 6B**), which is consistent with the

chloro substituted biphenyl-carbonyl modification of NS3. Compound **5** binds to bZiPro as the ^{1}H-^{15}N-HSQC of its complex with bZiPro exhibits obvious chemical shift difference from that of the free bZiPro (**Fig. 6C**). Superimposition of the ^{1}H-^{15}N-HSQC spectra of bZiPro in complex with compound **1** and compound **5** reveals profound chemical shift differences for numerous residues (**Fig. 6D**). The biphenyl group does not fit well in the S1 site based on the obtained structure of bZiPro-compound **1** complex. The observed chemical shift perturbations (CSPs) of several residues may be due to re-orientation of the biphenyl group in the active site upon covalent bond formation as modelled in **Fig. 6E**. As both compounds **1** and **5** are covalent inhibitors, IC_{50} values cannot accurately reflect the differences in inhibitory potential. To reveal the two inhibition potency matric of the two compounds, we then determined the specificity constants (k_{inact}/K_i). The Compound **5** with a k_{inact}/K_i value of 0.143 x 10^6 $M^{-1}s^{-1}$ is significantly more potent than compound **1** which has a k_{inact}/K_i value of 0.00458 x 10^6 $M^{-1}s^{-1}$ **(Fig. 7)**. The k_{inact} of compound **5** is about 36 times higher than that of compound **1** indicating that compound **5** is able to form adduct with and inactivate the protease at a much faster rate **(Fig.7)**.

Discussion

Several structures of flaviviral proteases in complex with peptidic inhibitors have been obtained while little structural information for protease in complex with small molecule inhibitors is available (Luo et al., 2015; Nitsche et al., 2014a). Free linked flaviviral proteases were shown to harbor both open conformation in which the β-hairpin region of the NS2B cofactor is dynamic and stays away from NS3 and closed conformation in which the β-hairpin region is engaged with NS3 and forms part of the active site (Erbel et al., 2006; Noble et al., 2012; Noble and Shi, 2012) (Figure S2). Our structural study has shown that the free unlinked construct-bZiPro exists in the closed conformation while conformational exchanges are also present in the absence of an inhibitor (Zhang et al., 2016). It has been well proven that ligand

binding can stabilize the closed conformation of proteases by forming interactions with residues from both NS3 and the C-terminal region of NS2B cofactor region (Su et al., 2009). In this study, we demonstrate that compound **1** is able to stabilize the closed conformation of bZiPro by suppressing the conformational exchanges (**Fig. 3**). Interestingly, compound **1** binding to bZiPro causes an increase in molecular mass of NS3 by 104 Da, which is due to the covalent modification of S135 by the benzoyl moiety (**Fig. 1, Fig. 2**). Based on the X-ray crystal structure, only part of S1 pocket is occupied by the benzoyl moiety (**Fig. 2**). The remaining hydroxyl-pyrazole derivative from compound **1** falls off the protease (**Fig. 2**). The integrity of compound **1** is essential for protease binding as the derived fragments from compound **1** exhibit no clear interactions with protease (Figure S6). Recently, we reported the crystal structure of bZiPro in complex with a fragment (EN300) (Zhang et al., 2016). This fragment interacts with part of the S1 pocket and overlap with the guanidinium group of P1 Arg, which is different from that of the benzoyl moiety (**Fig. 2** and **Fig. 3**). EN300 binding to bZiPro does not suppress the conformational exchanges as the missing cross peaks in the free bZiPro are still not observed in the bZiPro-EN300 complex (Zhang et al., 2016). It is interesting that modification of the side chain of S135 with a benzoyl group (M.W. 104 Da) can stabilize the closed conformation of the protease (**Fig. 2, Fig. 4**). Such result suggests that suppressing the conformational exchanges between NS2B and NS3 by an inhibitor does not require direct interactions between the inhibitor and NS2B.

We have confirmed that compound **1** and compound **5** exhibit micro to nano-molar IC_{50}s against bZiPro (**Fig. 1, Fig. 6 and Fig. 7**). Although compound **1** exhibits a weaker IC_{50} against bZiPro than WNV and DENV proteases, modification of compound **1** generates compound **5** which can also covalently modify bZiPro with much improved inhibition potency (**Fig. 6 and Fig. 7**). As compound **2** and compound **3** representing the corresponding fragments of compound **1** exhibit no obvious interactions with bZiPro (Figure S6), we speculate the following steps involved in inhibition of bZiPro by compounds **1**. The intact inhibitor first recognizes and binds to protease active site. Benzoylation of bZiPro on the side chain of S135 is the result

of the nucleophilic attack of S135 hydroxyl on the inhibitor ester, which occurs after compound binding. The formation of a covalent bond between the benzoyl moiety of compound **1** and the side chain of S135 stabilizes the closed conformation. The hydroxyl pyrazole group of compound **1** falls off from the protease. It is encouraging to identify this type of inhibitors because they are small molecules which might be able to overcome the cell penetrating problems that are encountered for substrate-derived inhibitors. It has also been noted that this type of inhibitors are not stable in solution because they can be hydrolyzed in buffer, which limits their further development. Nonetheless, our current study provides a piece of evidence that small molecular weight inhibitors can be developed against ZIKV protease while further effort is still needed due to the limitation of this type of inhibitors.

Structure-based drug design is a powerful strategy in target-based drug discovery. Knowing protein-ligand structural information is helpful for further inhibitor design. In this study, we characterized protein-inhibitor interactions using solution NMR spectroscopy, MS and X-ray crystallography. It is clear that the structures of NS2B and NS3 remain almost unchanged in the presence of different types of inhibitors such as fragment (Zhang et al., 2016), substrate (Phoo et al., 2016) or peptidic inhibitors (Lei et al., 2016; Li et al., 2017b). The close conformation of flavivirus protease should be used in structure-based drug design. Despite the stability issues of compound **1** in solution, current study provides structural information for bZiPro linked to a benzoyl moiety (**Fig. 2**). Our biochemical assay clearly demonstrates the inhibitory efforts of this type of inhibitors on protease activity, suggesting that it is possible to develop small molecule protease inhibitors. Compared with dengue and West Nile proteases, structure-based drug design may be a powerful tool to design potent inhibitors against Zika protease as numerous structures of the protease in the absence and presence of inhibitors are available(Kang et al., 2017) (Table S1).

In summary, we use X-ray crystallography and MS to demonstrate that compound **1** forms covalent interactions with bZiPro and the benzolated protease is in the closed conformation. The integrity of the

inhibitor is essential for the activity and protease binding. Modification of compound **1** generates a more potent inhibitor-compound **5**. Current study provides an example showing that the active site small molecule inhibitors are available while more chemical modifications might be required to improve the potency of the inhibitors.

STAR*METHODS

KEY RESOURCES TABLE

CONTACT FOR REAGENT AND RESOURCE SHARING

EXPERIMENTAL MODEL AND SUBJECT DETAILS

METHOD DETAILS

Protein expression and purification

bZiPro C143S site-directed mutagenesis

X-ray structure determination

NMR experiments

Mass spectrometry (MS) analysis

Inhibition assays

QUANTIFICATION AND STATISTIAL ANALYSIS

DATA AND SOFTWARE AVAILABILITY

ADDITIONAL RESOURCES

Supplemental information

Supplemental information includes Figures S1-S6 and Table S1.

AUTHOR CONTRIBUTIONS

Conceptualization, D.L. and C.K. Methodology, K.N., D.L., C.K., A.E.J. and J.H.. Y.R.L.; Investigation, Y.L., Z.Z., W.W.P., Y.R.L., H.Y.Y., R.L., K.N., A.E.J., Y.H., T.H.K., D.L., and C.K.; Writing-Original Draft, D.L. and C.K.; Writing-Review & Editing, D.L., C.K., K.N., J.H. and T.H.K.; Funding Acquisition, D.L. and C.K.; Resources, D.L., C.K., K.N., T.H.K., and J.H.; Supervision, D.L. and C.K.

Acknowledgements

We thank scientists from Australian Light Source MX beam-line and Swiss Light Source PX beam-line for their help with diffraction data collection. 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, (3) National Research Foundation grant NRF2016NRF-CRP001-063 to D.L. (4) A*STAR JCO grant (1431AFG102/1331A028) to C.K.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

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., *et al.* (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., and Adams, P.D. (2012). Towards automated crystallographic structure refinement with phenix.refine. Acta crystallographica. Section D, Biological crystallography *68*, 352-367.
Baronti, C., Piorkowski, G., Charrel, R.N., Boubis, L., Leparc-Goffart, I., and de Lamballerie, X. (2014). Complete Coding Sequence of Zika Virus from a French Polynesia Outbreak in 2013. Genome Announcements *2*.
Battye, T.G., Kontogiannis, L., Johnson, O., Powell, H.R., and 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., and 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., and Yang, H. (2016). Mechanisms of activation and inhibition of Zika virus NS2B-NS3 protease. Cell research *26*, 1260-1263.
Clum, S., Ebner, K.E., and Padmanabhan, R. (1997). Cotranslational membrane insertion of the serine proteinase precursor NS2B-NS3(Pro) of dengue virus type 2 is required for efficient in vitro processing and is mediated through the hydrophobic regions of NS2B. J Biol Chem *272*, 30715-30723.
Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J., and Bax, A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR *6*, 277-293.
Deng, J., Li, N., Liu, H., Zuo, Z., Liew, O.W., Xu, W., Chen, G., Tong, X., Tang, W., Zhu, J.*, et al.* (2012). Discovery of novel small molecule inhibitors of dengue viral NS2B-NS3 protease using virtual screening and scaffold hopping. J Med Chem *55*, 6278-6293.
Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr D *60*, 2126-2132.
Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallographica Section D *66*, 486-501.
Erbel, P., Schiering, N., D'Arcy, A., Renatus, M., Kroemer, M., Lim, S.P., Yin, Z., Keller, T.H., Vasudevan, S.G., and Hommel, U. (2006). Structural basis for the activation of flaviviral NS3 proteases from dengue and West Nile virus. Nat Struct Mol Biol *13*, 372-373.
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.
Gayen, S., Li, Q., Chen, A.S., Nguyen, T.H., Huang, Q., Hill, J., and Kang, C. (2011). An NMR study of the N-terminal domain of wild-type hERG and a T65P trafficking deficient hERG mutant. Proteins *79*, 2557-2565.
Headd, J.J., Echols, N., Afonine, P.V., Grosse-Kunstleve, R.W., Chen, V.B., Moriarty, N.W., Richardson, D.C., Richardson, J.S., and 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.
Johnston, P.A., Phillips, J., Shun, T.Y., Shinde, S., Lazo, J.S., Huryn, D.M., Myers, M.C., Ratnikov, B.I., Smith, J.W., Su, Y.*, et al.* (2007). HTS identifies novel and specific uncompetitive inhibitors of the two-component NS2B-NS3 proteinase of West Nile virus. Assay Drug Dev Technol *5*, 737-750.

Jones, C.L., Yeung, B.K., Manjunatha, U., Shi, P.Y., Bodenreider, C., and Diagana, T.T. (2015). Drug discovery for the developing world: progress at the Novartis Institute for Tropical Diseases. Nat Rev Drug Discov *14*, 442-444.
Kang, C., Keller, T.H., and Luo, D. (2017). Zika Virus Protease: An Antiviral Drug Target. Trends in microbiology *25*, 797-808.
Kay, L.E., Torchia, D.A., and Bax, A. (1989). Backbone dynamics of proteins as studied by 15N inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. Biochemistry *28*, 8972-8979.
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., *et al.* (2015). Identification of covalent active site inhibitors of dengue virus protease. Drug Des Devel Ther *9*, 6389-6399.
Lei, J., Hansen, G., Nitsche, C., Klein, C.D., Zhang, L., and Hilgenfeld, R. (2016). Crystal structure of Zika virus NS2B-NS3 protease in complex with a boronate inhibitor. Science *353*, 503-505.
Leung, D., Schroder, K., White, H., Fang, N.X., Stoermer, M.J., Abbenante, G., Martin, J.L., Young, P.R., and Fairlie, D.P. (2001). 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.
Li, Y., Li, Q., Wong, Y.L., Liew, L.S., and Kang, C. (2015). Membrane topology of NS2B of dengue virus revealed by NMR spectroscopy. Biochim Biophys Acta *1848*, 2244-2252.
Li, Y., Phoo, W.W., Loh, Y.R., Zhang, Z., Ng, E.Y., Wang, W., Keller, T.H., Luo, D., and Kang, C. (2017a). 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., *et al.* (2017b). Structural Dynamics of Zika Virus NS2B-NS3 Protease Binding to Dipeptide Inhibitors. Structure *25*, 1242-1250 e1243.
Lim, S.P., and Shi, P.Y. (2013). West Nile virus drug discovery. Viruses *5*, 2977-3006.
Lim, S.P., Wang, Q.Y., Noble, C.G., Chen, Y.L., Dong, H., Zou, B., Yokokawa, F., Nilar, S., Smith, P., Beer, D., *et al.* (2013). Ten years of dengue drug discovery: progress and prospects. Antiviral research *100*, 500-519.
Luo, D., Vasudevan, S.G., and Lescar, J. (2015). The flavivirus NS2B-NS3 protease-helicase as a target for antiviral drug development. Antiviral Res *118*, 148-158.
Ndeffo-Mbah, M.L., Parpia, A.S., and Galvani, A.P. (2016). Mitigating Prenatal Zika Virus Infection in the Americas. Annals of internal medicine *165*, 551-559.
Nitsche, C., Behnam, M.A., Steuer, C., and Klein, C.D. (2012). Retro peptide-hybrids as selective inhibitors of the Dengue virus NS2B-NS3 protease. Antiviral Res *94*, 72-79.
Nitsche, C., Holloway, S., Schirmeister, T., and Klein, C.D. (2014a). Biochemistry and medicinal chemistry of the dengue virus protease. Chemical reviews *114*, 11348-11381.
Nitsche, C., Holloway, S., Schirmeister, T., and Klein, C.D. (2014b). Biochemistry and Medicinal Chemistry of the Dengue Virus Protease. Chemical Reviews.
Noble, C.G., Seh, C.C., Chao, A.T., and Shi, P.Y. (2012). Ligand-bound structures of the dengue virus protease reveal the active conformation. J Virol *86*, 438-446.
Noble, C.G., and Shi, P.Y. (2012). Structural biology of dengue virus enzymes: Towards rational design of therapeutics. Antiviral Res *96*, 115-126.
Pervushin, K., Ono, A., Fernandez, C., Szyperski, T., Kainosho, M., and Wuthrich, K. (1998). 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.
Petersen, L.R., Jamieson, D.J., Powers, A.M., and Honein, M.A. (2016). Zika Virus. N Engl J Med *374*, 1552-1563.
Phoo, W.W., Li, Y., Zhang, Z., Lee, M.Y., Loh, Y.R., Tan, Y.B., Ng, E.Y., Lescar, J., Kang, C., and Luo, D. (2016). Structure of the NS2B-NS3 protease from Zika virus after self-cleavage. Nature communications *7*, 13410.

Potterton, E., Briggs, P., Turkenburg, M., and 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., and Keller, T.H. (2014). Drug design for flavivirus proteases: what are we missing? Curr Pharm Des *20*, 3422-3427.
PyMOL The PyMOL Molecular Graphics System, V., Schrödinger, LLC.
Salzmann, M., Pervushin, K., Wider, G., Senn, H., and Wuthrich, K. (1998). TROSY in triple-resonance experiments: new perspectives for sequential NMR assignment of large proteins. Proc Natl Acad Sci U S A *95*, 13585-13590.
Shen, Y., Delaglio, F., Cornilescu, G., and Bax, A. (2009). TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. J Biomol NMR *44*, 213-223.
Sidique, S., Shiryaev, S.A., Ratnikov, B.I., Herath, A., Su, Y., Strongin, A.Y., and Cosford, N.D. (2009). Structure-activity relationship and improved hydrolytic stability of pyrazole derivatives that are allosteric inhibitors of West Nile Virus NS2B-NS3 proteinase. Bioorg Med Chem Lett *19*, 5773-5777.
Su, X.C., Ozawa, K., Qi, R., Vasudevan, S.G., Lim, S.P., and 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.
Yang, C.C., Hsieh, Y.C., Lee, S.J., Wu, S.H., Liao, C.L., Tsao, C.H., Chao, Y.S., Chern, J.H., Wu, C.P., and Yueh, A. (2011). Novel dengue virus-specific NS2B/NS3 protease inhibitor, BP2109, discovered by a high-throughput screening assay. Antimicrobial agents and chemotherapy *55*, 229-238.
Yin, Z., Patel, S.J., Wang, W.L., Wang, G., Chan, W.L., Rao, K.R., Alam, J., Jeyaraj, D.A., Ngew, X., Patel, V., *et al.* (2006). Peptide inhibitors of Dengue virus NS3 protease. Part 1: Warhead. Bioorg Med Chem Lett *16*, 36-39.
Zhang, Z., Li, Y., Loh, Y.R., Phoo, W.W., Hung, A.W., Kang, C., and Luo, D. (2016). Crystal structure of unlinked NS2B-NS3 protease from Zika virus. Science *354*, 1597-1600.

Table 1 Data collection and refinement statistics

Data Collection Statistics	**bZiPro(C143S) + Compound1**	**bZiPro(C143S) + Ac-Lys-Arg-H**
Wavelength (Å)	0.9537	0.9537
Resolution range (Å)	45.74 – 1.90 (1.94 – 1.90) *	45.92 – 1.51 (1.54 – 1.51) *
Space group	$P2_122_1$	$P4_322$
Unit cell a, b, c, α, β, γ (Å) (°)	59.37, 59.63, 215.28 90.0, 90.0, 90.0	42.92, 42.92, 215.62 90.0, 90.0, 90.0
Total number of reflections	367855 (22198)	468463 (19329)
Unique reflections	61049 (3898)	33148 (1611)
Multiplicity	6.0 (5.7)	14.1 (12.0)
Completeness (%)	99.6 (100.0)	100.0 (100.0)

I/σI	9.9 (2.0)	14.6 (1.5)
Wilson B-factor (Å^2)	30.4	26.6
Rpim	0.058 (0.522)	0.025 (0.942)
CC½	0.998 (0.533)	0.997 (0.931)
$^aR_{merge}$	0.088 (0.749)	0.067 (2.178)
Refinement Statistics		
Resolution range (Å)	42.07 – 1.90(1.81- 1.75)	42.92 – 1.51(1.56- 1.51)
$^bR_{work}$ (%)	23.59	18.78
$^cR_{free}$ (%)	27.14	20.16
Number of non-hydrogen atoms	5942	1571
Macromolecules	5684	1473
Ligands	32	23
Water	226	98
Protein residues	755	195
RMSD (bonds) (Å)	0.005	0.006
^{d}RMSD (angles) (°)	1.05	1.20
Ramachandran favoured (%)	94.24	96.89
Ramachandran allowed (%)	5.49	3.11
Ramachandran outliers (%)	0.27	0.00
Clashscore	13.61	2.79
Average B-factor (Å^2)	29.4	42.1
Macromolecules	29.3	41.7
Ligands	24.7	42.3
Solvent	34.6	48.2

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

$^aR_{merge} = \Sigma_h \Sigma_i |I_{hi} - \langle I_h \rangle| / \Sigma_{h,i} I_{hi}$, where I_{hi} is the ith observation of the reflection h, while $\langle I_h \rangle$ is its mean intensity.

^{b} Rfactor $= \Sigma\, ||F_{obs}| - |F_{calc}|| / \Sigma |F_{obs}|$.

^{c} Rfree was calculated with 5% of reflections excluded from the whole refinement procedure.

Figure legends

Figure 1. Compound **1** forms covalent interactions with bZiPro. A. Chemical structure of compound **1**. B. MS indicates the covalent modification of bZiPro by compound **1**. The increase of the molecular weight by 104 Da suggests that protease is modified by the benzoyl group. C. Chemical structures of possible hydrolysis products of compound **1**. Compound **1** can be hydrolyzed due to the presence of ester group and the resulting products include a hydroxyl pyrazole moiety-compound **2**.

Figure 2. Structure of bZiPro bound to compound 1 and other inhibitors. A. Crystal structure of bZiPro-compound **1** complex. Compound **1** forms hydrogen bond with H51 and π-π stacking interactions with Y161 as marked green. B. The mFo-DFc map (green) contoured at 3σ and the 2mFo-DFc map (blue) at 1σ are shown for compound **1** and the active sites of protease. Potential hydrogen bonds are shown by the black dashed lines. NS2B protein residues are shown in light magenta and NS3 protein residues are shown in yellow. C. Structure of bZiPro-Ac-Lys-Arg-H complex. D. Structure of bZiPro-EN300 complex. EN300 is shown in cyan.

Figure 3.Comparison of Ac-Lys-Arg-H with wild type and C143S mutant of bZiPro in active sites.

A. The *mFo-DFc* map (green) contoured at 3σ and the *2mFo-DFc* map (blue) at 1σ are shown for Ac-Lys-Arg-H and side chain of the catalytic S135 . B. Superposition of bZiProC143S-Ac-Lys-Arg-H and bZiPro-Ac-Lys-Arg-H (PDB ID:5H6V). C. Detailed interactions of bZiProC143S-Ac-Lys-Arg-H complex in active sites. Ac-Lys-Arg-H shown in orange and protein residues shown in yellow. The electron density map (*2mFo-DFc*;3σ) of protein residues($NS3_{132-134}$) as shown in light blue. D. Detailed interactions of bZiPro – Ac-Lys-Arg-H complex in active sites. Ac-Lys-Arg-H shown in green and protein residues shown in pink.

Figure 4. Compound **1** binds to bZiPro in solution. A. Superimposed ^{1}H-^{15}N-HSQC spectra of bZipro in the absence (black) and presence (red) of compound **1**. ^{1}H-^{15}N-HSQC spectra of ^{15}N, ^{13}C, ^{2}H-labeled bZiPro (0.8 mM) in the absence ad presence of 1.6 mM of compound **1** were collected and processed. B. Residues of bZiPro affected by compound **1** binding. The structure of free bZiPro (PDB id 5GPI) is shown. Residues with observed cross peaks in the spectrum upon inhibitor binding are highlighted in green. Residues with averaged chemical shift perturbation (CSP) more than 0.12 ppm are shown in spheres and highlighted in red. Other residues from NS2B and NS3 are shown in purple and yellow, respectively. H51, D75 and S135 are labeled. The figures are made using PyMOL. C. Titration of compound **1** to bZiPro. The ^{1}H-^{15}N-HSQC spectra of bZiPro in the absence (black) and presence of different amounts of ligands are superimposed.

D. Several residues undergo slow exchanges. Several residues such as L78 and L140 exhibit multiple peaks when protein to ligand ratio is more than 1. E. Several residues exhibit cross peaks in the spectra when bZiPro is mixed with compound **1**.

Figure 5. Structure and dynamics of bZiPro-compound **1** complex. A. The ^{15}N-R_1, R_2 and hetNOEs values are plotted against residue number. The data were collected on a Bruker Avance 600 MHz magnet equipped with a cryoprobe. A sample containing 0.8 mM bZiPro and 1.6 mM compound **1** was used for data acquisition. The data were collected at 25 °C and the values were obtained using NMRView. B. Structure of the bZiPro-compound **1** complex. The structure is obtained in this study and the color codes are same as Fig. 4. Residues with identified NOEs are shown in blue spheres. C. NOESY spectroscopy of the complex. Strip plots of the NOEs of residues are shown. Lines indicate the unambiguously identified NOEs between amide protons.

Figures 6. Compound **5** binds to bZiPro. A. IC_{50} of compound **5** against bZiPro. Inset is the structure of compound **5**. B. compound **5** binds covalently to bZiPro. MS spectra indicate a 204 Da increase of molecular mass of NS3 upon compound binding, suggesting the ester group is not stable in solution. C. ^{1}H-^{15}N-HSQC spectra of bZiPro in the absence (black) and presence (red) of compound **5**. D. superimposed ^{1}H-^{15}N HSQC spectra of bZiPro in the presence of compounds **1** and **5**. E. Models of the bZiPro-compound **5** complex.

Figure 7. The k_{inact}/K_i values of Compounds **1** and **5** on bZiPro. (A)(B) The time dependent inhibition of bZiPro at different concentrations of compound **1** (A) and compound **5** (B). (C)(D) The k_{obs} of the two compounds were plotted against inhibitor concentration as described in the method section. The k_{inact}/K_i value of individual compounds was determined for compound **1** (C), and compound **5** (D). E. The k_{inact}/K_i values for different compounds.

STAR*METHODS

KEY RESOURCES TABLE

Reagent or resource	Source	Identifier
Compounds and vectors		
Compounds **1-5**	Synthesized in this study	N/A

Bz-NKRR-AMC	Peptides International	Cat No.:MCA-3923-PI
pETDUET-bZiPro	(Phoo et al., 2016; Zhang et al., 2016)	N/A
Protein purification		
Ni^{2+}-NTA resin	Qiagen	Cat No./ID: 30230
HiPrep™ 16/60 Sephacryl™-S200 HR column	GE Healthcare	Code:17-1166-01
PD-10 desalting columns	GE Healthcare	Code: 17-0851-01
Amicon® Ultra-15 (MWCO 10 kDa)	Merck	Ref:UFC901096
Chemicals used for making media and buffers		
$^{15}NH_4Cl$	Cambridge Isotope Laboratory	Cat no.: NLM-467
^{13}C-glucose	Cambridge Isotope Laboratory	Cat no.: CLM-1396
Deuterium Oxide (99%)	Cambridge Isotope Laboratory	Cat no.: DLM-4-99
HEPES	1st Base	C/No: BUF-1821
Tris-HCl	1st BASE	C/No:BUF-1415
EDTA	1st BASE	C/No:BUF-1053
NaCl	Merck	Cat No.: 1.06404
Na_2HPO_4	Sigma-Aldrich	Cat No.: S5136
NaH_2PO_4	Sigma-Aldrich	Cat No.: S9638
Dithiothreitol (D)	Goldbio	Cat No.: DTT25
β-mercapthoethanol	Sigma-Aldrich	Cat No.: M6250
β-D-1-thiogalactopyranoside	Goldbio	Cat No.: I2481C

Dimethyl sulfoxide (DMSO)	Sigma	Cat: D2650
imidazole	Sigma	Cat: P9791
Glycerol	Sigma	Cat: g5516
Protein Analysis		
SeeBlue^{R} Plus molecular weight marker	Invitrogen	Cat: LC5925
Running buffer	Invitrogen	
NuPAGE NOVEX 4-12% Bis-Tris gels	Invitrogen	Ref:NP0323BOX

CONTACT FOR REAGENT AND RESOURCE SHARING

Plasmids and other reagents generated in this study or any other question about the reagents please contact luodahai@ntu.edu.sg or CongBao Kang, cbkang@etc.a-star.edu.sg.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

METHOD DETAILS

Materials and methods

Protein expression and purification

The 1-177 of NS3 of ZIKV (GenBank: HQ234500.1) containing an N-terminal His6-tag and a thrombin was co-expressed with a fragment comprising amino acids 45-96 of NS2B in *E. coli* strain BL21 (DE3) and Rosetta (DE3) under control of T7 promotors. The transmembrane domains of NS2B and the protease recognition sequence were not included. The bacterial cultures were grown in M9 medium or LB broth at 37 °C and 200 rpm to an optical density (OD_{600}) of 0.8. Protein expression was induced overnight at 18 °C by adding β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 1 mM. Cells were harvested by centrifugation at 10,000 ×g for 10 min and the resulting pellets were stored at -80 °C. ^{15}N, or ^{13}C, ^{15}N labeling was achieved by using 1 g/L $^{15}NH_4Cl$, 2 g/L ^{13}C-glucose as the sole nitrogen and carbon sources. D_2O was used to replace the water in the M9 medium when deuteration was required. For expression protein in D_2O, bacterial were stepwise-adapted in two different M9 media with 85% and 100% D_2O before protein expression was conducted.

Cell pellets were re-suspended in a lysis buffer (20 mM sodium phosphate, pH 7.8, 500 mM NaCl, 2 mM β-mercaptoethanol, 10 mM imidazole and 5% glycerol). Cells were broken using a LM20 microfluidizer (Microfluidics, USA) or a Q700-220 sonicator (Misonix, USA). Lysate was cleared by centrifugation at 40, 000 ×g and 4 °C for 20 min. Recombinant protein was bound to Ni^{2+}-NTA beads (Qiagen, USA) that were preequilibrated with the lysis buffer. After washing with a washing buffer (20 mM sodium phosphate, pH 7.8, 1 M NaCl, 2 mM β-mercaptoethanol, 20 mM imidazole and 5% glycerol), protein was eluted with the elution buffer (500 mM imidazole, pH6.5, 500 mM NaCl, 2 mM β-mercaptoethanol and 5% glycerol). Protein was finally subjected to gel filtration on a 16/600 Superdex 200 column (GE Healthcare, USA equilibrated with 20 mM HEPES pH 7.5, 150 mM NaCl, 5% glycerol, and 2mM DTT or NMR buffer (20 mM HEPES pH 7.3, 150 mM NaCl, and 2mM DTT). For crystallization, thrombin was added to the sample to remove the N-terminal tag before the gel filtration step. Protein was concentrated using centrifugal concentrators (Vivaspin, Satorius, 10 kDa MWCO) and stored at -80 °C.

bZiPro C143S site-directed mutagenesis

The bZiPro variant was produced by using synthesized codon optimized cDNA sequence of NS2B-NS3 (Genescript) as template. Two separate fragments as 45-96 of NS2B and 1-177 of NS3 were selected with Cys143Ser mutation site designed in primers. The Forward primer used as GGATCTCCGATCCTCGACAAGTCAGGACGTGTGATCGGACTTTATGGC, and Reverse as GCCATAAAGTCCGATCACACGTCCTGACTTGTCGAGGATCGGAGATCC. The fragment was amplified by KODFX Neo of TOYOBO. The PCR product was transformed to DH5 α. The positive clones were selected and plasmids were extracted after overnight culture. The plasmid with Cys143Ser mutation was verified by sequencing and expressed for further purification.

X-ray structure determination

Compound **1** was dissolved in dimethyl sulfoxide to a 100 mM stock solution. The bZiProC143S protein crystallization was determined as described (Zhang et al., 2016). The compound **1** was added to protein (molar ratio 3:1) at the concentration of 1 mg/ml and concentrated to 40 mg/ml for hanging vapor diffusion method. The bZiPro - compound **1** complex was mixed with reservoir solution (0.2 M Ammonium sulfate, 0.1 M Sodium acetate trihydrate pH 4.6, 25% PEG4000) as 1:1 in volume ratio. Totally 2 µL was incubated at 18 ^{o}C until crystals appeared. Then the crystals were selected and soaked in protect solution (0.2 M Ammonium sulfate, 0.1 M Sodium acetate trihydrate pH 4.6, 25% PEG4000, 20% glycerol), further flash frozen in liquid N_2.

Diffraction data was collected at Australian Synchrotron (Clayton, Victoria) by using the MX1 beamline. IMOSFLM Software was used for diffraction intensities integration (Battye et al., 2011). Data collection statistics was analyzed using POINTLESS and AIMLESS function of CCP4 suite (Evans, 2007, 2011; Potterton et al., 2003), as listed in Table 1. Refinement was proceed by Phenix (Adams et al., 2010; Afonine et al., 2012; Headd et al., 2012)and Coot (Emsley and Cowtan, 2004; Emsley et al., 2010). PyMOL was used for generate figures (PyMOL The PyMOL Molecular Graphics System).

NMR experiments

Isotope-labeled bZiPro were typically concentrated to 0.8 to 1.0 mM in a buffer that contained 20 mM HEPES pH7.3, 150 mM NaCl and 1 mM DTT. All NMR experiments were performed at 298 K on a Bruker Avance II 700-MHz spectrometer. Sequential assignment was obtained using transverse relaxation-optimized spectroscopy (TROSY) (Pervushin et al., 1998; Salzmann et al., 1998)-based experiments including 2D-^{1}H-^{15}N-HSQC, 3D-HNCACB, 3D-HNCOCACB, 3D-HNCA, 3D-HNCO and NOESY-TROSY (with a mixing time of 120 ms). To identify the NOEs between protein and compound **1**, an F1-^{13}C/^{15}N-filtered and F2-^{15}N-edited NOESY experiment (mixing time 200 ms) was collected using a ^{15}N, ^{13}C, and ^{2}H-labeled bZiPro in complex with compound **1**. All the pulse programs used in this study are from the standard pulse library (Topspin 2.1, Bruker). The NMR data were processed with the NMRPipe (Delaglio et al., 1995) and Topspin 2.1. Data were visualized with NMRView (Johnson, 2004) and Cara. Protein secondary structures of the complex were determined using TALOSN (Shen et al., 2009). The ^{15}N relaxation parameters including ^{15}N- R_1, R_2 and hetNOE (Kay et al., 1989) were obtained from the standard relaxation experiments.

Mass spectrometry (MS) analysis

Modification of protease by the compound was analyzed using MS as previously described (Koh-Stenta et al., 2015). Purified bZiPro was diluted to 20 µM using the NMR buffer. Compound from the stock was mixed with bZiPro with a compound-protein ration of 4:1 at room temperature for 1 h. DMSO was used as the reference. Samples were subjected to analysis and the molecular weight of protease was determined. Two major peaks corresponding to NS2B cofactor region and NS3 protease region were observed in the spectra because there is no linker present in these two polypeptides. Increase of the molecular weight of NS3 indicates the covalent modification by the compound. S135 of NS3 is responsible for the covalent linage with compound **1**, verified by mutagenesis study (Koh-Stenta et al., 2015).

Inhibition assays

Inhibition assay was carried out using the method described previously (Phoo et al., 2016). Briefly, synthesized inhibitors were dissolved in DMSO. The assay buffer contained 20 mM Tris pH 8.5, 10% glycerol, 0.01% Triton X-100 (Leung et al., 2001). Protease was diluted to 3 nM and incubated with different amounts of inhibitors in the assay buffer for 30 min at a final DMSO concentration of 0.3%. Substrate Bz-nKRR-AMC was added to a final concentration of 30 µM. The fluorescence intensities of the mixtures were monitored at 60 second intervals over 10 min using a Cytation 3 Mulitmode plate reader (BioTek). The reaction and monitor temperature was at 37 °C. The excitation and emission wavelengths were 380 nm and 460 nm, respectively. All the assays were carried out in Corning® 96-well plates as triplicates. The half maximal inhibitory concentration (IC_{50}) of an inhibitor was determined using GraphPad Prism version 5.0 (GraphPad Software, La Jolla California, USA).

To determine the k_{inact}/K_i of compound **1** and compound **5**, the different concentrations of the inhibitor compound are mixed with the Bz-NKRR-AMC substrate at 30 µM. The reaction was started by addition of bZiPro. The final conc. of bZiPro was 1 nM. The rates were monitored over 1 hour, recording the fluorescent readings at 2 minutes intervals at excitation and emission wavelengths of 380 nm and 460 nm respectively. The recorded fluorescent intensities are plotted against time using pseudo-first-order equation in GraphPad Prism version 5.0 to obtain the k_{obs}. The k_{inact}/K_i was obtained by plotting the k_{obs} against the inhibitor concentrations using the function one-site-total-binding in GraphPad Prism version 5.0.

QUANTIFICATION AND STATISTICAL ANALYSIS

^{15}N relaxation data collection and analysis

^{15}N relaxation parameters including ^{15}N- R_1 ($1/T_1$), R_2 ($1/T_2$) and hetNOE (Kay et al., 1989) were collected at 25 °C for aforementioned sample as we described previously (Li et al., 2017b). For the R_1 measurement, several ^1H-^{15}N-HSQC experiments with different relaxation delays of 100, 200, 400, 600, 800, 1200, 1600, 2000, 2500 and 3000 ms were collected and analyzed, respectively. For the R_2 measurement, ^1H-^{15}N-HSQC spectra with relaxation delays of 17, 34, 51, 68, 85, 102, 119, 136 and 153 ms were obtained and analyzed, respectively. The hetNOEs were obtained using two datasets that were collected with and without initial proton saturation for a period of 3 s (Gayen et al., 2011). The R_1, R_2 and hetNOE values were calculated using NMRView (Johnson, 2004). The titration experiments were analyzed and the binding constants were obtained using NMRView Java version (Johnson, 2004).

DATA AND SOFTWARE AVAILABILITY

Coordinates and structure factors have been deposited in the Protein Data Bank under accession code 5YOF and 5YOD. Assignment of protease-compound **1** complex has been deposited in the Biological Magnetic Resonance Data Bank (BMRB) with accession number 27282.

Softwares

NMRVIEWJ	Johnson, 2004	http://www.onemoonscientific.com/nmrviewj
NMRPipe	Frank Delaglio, 2016	https://www.ibbr.umd.edu/nmrpipe/install.html
CARA	Kurt Wüthrich, 2015	http://wiki.cara.nmr.ch/InstallCara
Topspin (2.1, 3.5)	Bruker	https://www.bruker.com/products/mr/nmr/nmr-software/software/topspin/overview.html
Sparky	T. D. Goddard and D. G. Kneller, 2008	https://www.cgl.ucsf.edu/home/sparky/
PyMOL (1.8)	Schrödinger, 2016	http://pymol.org/support/platforms

Prism	Graphpad 6, 2015	https://www.graphpad.com/scientific-software/prism/
iMOSFLM	Harry Powell, 2015	http://www.mrc-lmb.cam.ac.uk/harry/imosflm/ver721/introduction.html
AIMLESS	Phil Evans, 2013	http://www.mrc-lmb.cam.ac.uk/harry/pre/aimless.html
MOLREP	A.A.Vagin, 2010	http://www.ccp4.ac.uk/html/molrep.html
Phenix	P. D. Adams, 2010	https://www.phenix-online.org/
Coot	P. Emsley, 2010	https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/