Zika Virus Protease: An Antiviral Drug Target

CongBao Kang[1*], Thomas H. Keller[1*], and Dahai Luo[2,3*]

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.

*Correspondence: cbkang@etc.a-star.edu.sg (C. Kang); thkeller@etc.a-star.edu.sg (T.H. Keller); luodahai@ntu.edu.sg (D. Luo).

Key words: flavivirus; protease Inhibitor; structure-based drug design.

Abstract

The recent outbreak of Zika virus (ZIKV) infection has caused global concern due to its link to severe damage to the brain development of foetuses and neuronal complications in adult patients. A worldwide research effort has been undertaken to identify effective and safe treatment and vaccination options. Among the proposed viral and host components, the viral NS2B-NS3 protease represents an attractive drug target due to its essential role in the virus life cycle. Here, we outline recent progress in studies on the Zika protease. Biochemical, biophysical and structural studies on different protease constructs provide new insight into the structure and activity of the protease. The unlinked construct displays higher enzymatic activity and better mimics the native state of the enzyme and therefore is better suited for drug discovery. Furthermore, the structure of the free enzyme adopts a closed conformation and a preformed active site. The availability of biophysical data for a fragment hit and peptide inhibitors, as well as the attainability of soakable crystals suggest that the unlinked construct is a promising tool for drug discovery.

Zika Virus

Zika virus (ZIKV) is a member of mosquito-borne flavivirus genus which contains important human pathogens such as dengue virus (DENV), West Nile virus (WNV), Japanese encephalitis virus (JEV) and yellow fever virus. ZIKV came to the attention of the general public in 2015, when a large outbreak occurred in Brazil and rapidly spread to other countries in the region [1]. This particular flavivirus had been known since 1947, but was considered fairly benign causing mild disease in about 20% of infected people with symptoms including fever, malaise and sometimes a rash and/or conjunctivitis [2]. The first public health concerns emerged in 2007 when an outbreak on the Micronesian island of Yap showed that neurological disorders seemed to increase during a Zika virus outbreak [3].

During the 2015 epidemic in South America, clusters of microcephaly and neurological disorders were reported [4] and tentatively linked to the circulation of the ZIKV, prompting the Director-General of WHO to declare a Public Health Emergency of International Concern [2]. Since then the association between laboratory-confirmed Zika virus infections and microcephaly has been strengthened [5] but still further epidemiological evidence is needed to finally prove the connection.

It is currently unclear what turned a relatively benign virus into a pandemic threat, but experts have been pointing out that immune enhancement by pre-existing heterologous anti-flavivirus antibodies may be involved in the complications associated with ZIKV infections [4]. Evidence that supports this hypothesis was recently reported by Bardina et al. who showed that mice with ZIKV infections receiving Dengue-positive donor plasma showed increased morbidity and mortality [6]. Given the complications of ZIKV infection and the potential damage to the public heath, effective treatment options including vaccines and pharmacological interventions are pursued by academia and industry all around the world.

Vaccine and Drug Development

Large outbreaks of viral infections invariably lead to discussions about the best medical countermeasures. Prophylactic vaccines have successfully been developed against yellow fever virus and Japanese encephalitis virus infections, and Institute Butantan (Brazil) has announced a collaboration with the NIH for an expedited Zika vaccine development program [7]. Several other Zika vaccine candidates are advancing towards clinical development [8]. Unfortunately the development of such a vaccine might be complicated by the well-known tendency of the flavivirus genus for antibody-mediated enhancement of infection. As a consequence, such a vaccine will require a careful evaluation of safety parameters and risk-benefit analysis for the affected region, potentially leading to extended and costly clinical trials [7].

Antiviral drugs have successfully been used to treat hepatitis C virus (HCV) infections and screening of existing antiviral libraries might lead to a rapid breakthrough for an antiviral to treat ZIKV [9, 10]. Development of a specific Zika-antiviral will however take several years since a drug candidate will have to pass through the standard clinical development. In this review we will discuss the prospects for a ZIKV NS2B-NS3 protease inhibitor and the possibility of developing a broad-spectrum flavivirus antiviral, which might be especially attractive in the current situation where dengue or West Nile infections may be responsible for the complications seen in the recent ZIKV outbreak.

Flaviviral NS2B-NS3 Protease

The Zika virus genome contains a protease that is essential for the replication of the virus. This protease performs the posttranslational processing of the polyprotein into the three structural (C, prM, and E) and the seven non-structural proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5) that are required to complete the virus life cycle and produce progeny viral particles (**Fig. 1A, Key Figure**). For all the members of the flavivirus genus, this processing is conserved and performed by the viral and host proteases. The chymotrypsin-like viral protease is made up of two components, namely the membrane protein NS2B and the N-terminal domain of NS3 which harbours the catalytic triad (H51, D75, and S135)(**Fig. 1B**). The central cofactor region

(approximately 40 residues) of NS2B is required for both correct folding and catalytic activity of the protease.

Although the natural form of the dengue protease harbouring full-length NS2B and the N-terminal region of NS3 has been expressed and purified from *Escherichia coli*, the presence of detergent micelles in the system limits its application in structural studies and inhibitor screening [11, 12]. The non-covalent interactions of these two viral proteins to form the all-important active site have posed a challenge for flavivirus drug discovery, since in vitro assays and structural studies are usually performed without a membrane component leading to uncertainty on how to mimic the natural conformation of the enzyme complex (**Box 1**).

For DENV, WNV, and JEV viral proteases, the structures were determined using a single-chain construct with the NS2B cofactor region linked to the NS3 protease region through a glycine-rich ($Gly_4SerGly_4$) linker (Reviewed recently [13, 14]). These studies identified a dual function for the NS2B cofactor region, stabilizing the correct protein fold of the core structure via NS2B 45-67 as a beta-strand and forming part of the substrate binding pocket (S2 pocket) via NS2B 68-96 beta-hairpin structure [15-18] (**Fig. 1**). Given the high sequence identity (~30-80%), the structures of these flaviviral proteases are very similar. Generally, the protease adopts two major conformations: the open conformation of the enzyme where the NS2B cofactor region is disordered and the active site is not completely formed; or the closed conformation where the NS2B cofactor is fully engaged with the NS3 protease for inhibitor recognition at the P2 position, allowing catalytic activity of the enzyme (**Fig. 1 and 2**).

Consideration of Various Constructs for Inhibitor Discovery

NMR studies suggest that a dengue virus protease complex formed by co-expression of NS2B and the NS3 protease domain (without a linker) is more stable and predominantly existing in the closed conformation, even in the absence of a ligand[19]. The importance of these findings for drug discovery is still controversial. However, for the dengue virus NS2B-NS3 protease it has never been possible to obtain multiple inhibitor-protein

complexes for computer aided design, a strategy that was successfully used for HCV protease drug discovery. The active protease has been produced in a bacterial expression system and the enzyme's specificity for synthesized FRET-type substrate libraries has been profiled [20, 21].

Three ZIKV protease constructs have been reported which exhibited different enzymatic kinetics (**Box 1**). The unlinked construct bZiPro has slightly higher K_M value and five times higher k_{cat} value (5.30 ± 0.35 s^{-1}) than the linked one gZiPro (1.07 ± 0.05 s^{-1}) and is a more efficient enzyme[22]. A recent biochemical study also shows that linked and unlinked proteases exhibit different enzymatic activities [23]. In this study, the unlinked protease was slightly different and likely the result of autoproteolysis at the junction of NS2B cofactor and G_4SG_4 linker of the linked construct. The unlinked protease has a K_M value of 49.25 ± 1.91 µM and a kcat value of 0.95 ± 0.01 s^{-1} whereas the linked has a K_M value of 61.03 ± 0.18 µM and a kcat value of 0.20 ± 0.02 s^{-1} [23]. The differences in the absolute values of the kinetic parameters between the two studies are mainly due to different substrates used as Phoo et al. used Bz-nKKR-AMC whereas Kuiper et al. used Pyr-TGKR-AMC [22, 23].In both cases, the unlinked protease has a higher kcat value and is more efficient than the linked protease, which also suggests that the protease is sensitive to the artificial linker between NS2B and NS3. Although the gZiPro is active in the enzymatic assays, the flexible linker may affect the access of the inhibitor to the active site and therefore interfere with the catalysis (**Fig. 1**). It has also been demonstrated recently that the gZiPro tends to favour the open conformation in solution [24]. The bZiPro construct would be more preferred for structure based antiviral inhibitor discovery for the following reasons: The structures were determined to very high resolution; The closed conformation offers the ready-to-bind active site which can be readily soacked or co-crystalized with inhibitors; Complex structures of bZiPro with a small fragment compound, peptide and reverse peptide inhibitors are available; The protein also displays very well resolved NMR profiles and majority of the residues have been assigned in the ^{1}H-^{15}N- heteronuclear single quantum coherence spectroscopy (HSQC) spectrum, making solution NMR based screening straightforward (**Fig. 1**) [22, 25, 26].

Membrane associated NS2B-NS3 is another in vitro construct for inhibitor screening and characterization (**Fig. 1F**). The relative orientation of the protease to the membrane may impact inhibitor access. The C-terminal tail of NS2B may also interact (and interfere) with the protease post NS2B-NS3 junction cleavage. Similar protease constructs have been reported for dengue protease [19, 27, 28]. It was also noticed that there is a conserved NS2B-NS3 dimer interface which is mediated by a set of polar contacts from both NS2B cofactor and NS3 [22, 29, 30]. The dimerization may be further driven by the membrane association of the polyprotein in the in vivo membrane-rich environment where the effective local concentration of the replicative enzymes is very high. Indeed, full length trans-membrane protein NS2B from dengue virus dimerizes in a cell-based assay [31]. However, the biological significance of the dimerization has yet to be demonstrated. Cross-validation with these different protease constructs may help remove false positive hits in compound screening campaigns and map the binding mode in the absence of a complex structure.

Crystal Structures of ZIKV NS2B-NS3 Protease

Several groups have reported structures of the protease as a free enzyme and inhibitor-bound complexes (**Table 1**, **Fig. 1 and 2**) [22, 25, 26, 29, 30, 32]. Solution NMR studies revealed additional information about the structural dynamics of the enzyme upon ligand binding [24-26, 29, 32]. To this end, three constructs - gZiPro, eZiPro, and bZiPro - have been studied enzymatically and structurally (**Fig. 1B and Box 1**). Similar to the earlier flaviviral linked NS2B-NS3 proteases, gZiPro has been captured in a closed conformation in the presence of a covalently bound boronate dipeptide inhibitor (**Fig. 1E**)[30]. In the free gZiPro, the C terminal half of the NS2B cofactor is disordered and the associated regions on NS3 are also not folded properly. The overall folding of the enzyme remains an open conformation as a large number of residues are missing in the structure (**Fig. 1F**) [29, 32]. Overall, the open and closed conformations of gZiPro are very similar to other flaviviral proteases of the same construct design and refreshed our memory of the earlier work done on other flaviviruses [13, 14].

eZiPro was designed to mimic the native form of NS2B-NS3, which is achieved by removing the NS2B trans-membrane regions and remaining a native NS2B-NS3 cleavage site (**Fig. 1B and Box 1**). As a result, the protein cleaves itself during sample preparation. The crystal structure of fully cleaved eZiPro reveals a closed conformation where the last four residues of NS2B occupy the protease active site while the NS3 N terminal residues are released (**Fig. 1D**). This structure captures for the first time a post-proteolytic state of the enzyme during viral polyprotein processing, reveals the interactions made by the carboxylic end of NS2B with the protease active site and suggests an auto-inhibitory role for the NS2B C-terminal peptide at the NS2B-NS3 junction following self-cleavage. The requirement of positively charged residues for the P1 and P2 sites is evident (**Fig 2A**). One interesting difference between ZIKV protease and the protease from other flaviviruses such as dengue virus or West Nile virus is the presence of Gly128 in the TGKR peptide (a Lysine in dengue NS2B-NS3 protease). Gly128 incompletely fills the S3 pocket and this structural information can be used to design tighter specific inhibitors [22].

To better mimic the protease function *in situ,* the un-linked protease (bZiPro) was used for structural studies. Interestingly, the structure of bZiPro was captured in a closed conformation, which appears to contain a preformed stable substrate binding pocket (**Fig. 1C**). [26]. It was also demonstrated that this closed conformation does not undergo further significant conformational changes upon substrate or inhibitor binding. It is worth noting that the crystal structure of a closed conformation of a free enzyme without any bound inhibitor has never been reported before for any flaviviruses. Previous structural studies of the flaviviral proteases suggest that substrates (or competitive inhibitors) induce conformational changes of NS2B cofactor (residues 62-96) to fold closely to the active site and participate in ligand binding. However previously reported protease crystal structures were all determined with the linked construct. Although being flexible and disordered in all structures, the glycine-rich linker may cause some steric hindrance or interference for the enzyme to fold properly while the linked protease remains stable in the presence of a strong inhibitor [13, 14].

Protein Dynamics of Zika NS2B-NS3 Protease in Solution

Solution NMR studies have been carried out to understand the structure and dynamics of flaviviral proteases in solution. The first study was performed on West Nile protease, in which the protease was shown to exist predominantly in the closed conformation in solution using NMR spectroscopy [33]. As the chemical shifts of the amide proton and the amides are sensitive to the environment surrounding the residues, the conformational exchanges can be monitored using ^{1}H-^{15}N-HSQC. Dissociation exchanges between NS2B and NS3 occur in West Nile protease, which leads to line broadening of cross peaks in the^{1}H-^{15}N-HSQC spectrum [33]. Further studies were carried out for dengue protease and the open conformation was found to be dominant in the linked protease [34]. Upon binding to an inhibitor, the closed conformation in both West Nile and dengue proteases is stabilized, which gives rise to nicely dispersed cross peaks in the spectra. Using a co-expression system, the unlinked dengue protease complex was produced [19]. This novel protease construct was confirmed to be in the closed conformation in solution and exhibited nicely dispersed cross peaks in the ^{1}H-^{15}N-HSQC spectrum, giving rise to the prediction that the artificial linker might affect protein dynamics to alter the populations of the open and closed conformations [19].

Based on the available knowledge, the three distinct Zika protease constructs were made and characterized by NMR spectroscopy [22]. The three constructs exhibited similar patterns in the ^{1}H-^{15}N-HSQC spectra. While the data suggests that the protease from all constructs have similar folding (**Fig. 3**), small differences in the NMR spectra provide further insights. Especially, chemical shifts differences observed between gZiPro and bZiPro suggest that the glycine-rich linker affects the environment of several residues through direct interactions and allosteric effects [22]. A binding experiment with a small protein bovine pancreatic trypsin inhibitor (BPTI) also suggests that the linker could affect large-molecular weight inhibitors entering the active site. Although the linked protease construct (gZiPro) exhibited measurable enzymatic activities, NMR binding studies reveal that the linked protease does not appear to bind well to BPTI - a potent inhibitor against both West Nile and dengue proteases [22, 35]. NMR studies on the unlinked Zika protease (bZiPro) provide complementary information

to the X-ray structures. The unlinked protease has an open active site and exhibits obvious interaction with a weak peptidic inhibitor (IC_{50}>100 μM, Kd in μM range), after which the protease conformation is stabilized to generate a well-defined ^{1}H-^{15}N-HSQC spectrum (**Fig. 3**). The Zika protease also forms a complex with the TGKR peptide verified by both NMR studies and crystal structure [22]. In addition, NMR studies have shown that unlinked Zika protease binds to a small chemical fragment EN300. Unlike peptidic inhibitors, fragment binding does not caused significant changes in the ^{1}H-^{15}N-HSQC spectrum because it binds only part of the S1 pocket and is away from the NS2B-NS3 binding interface [26]. The linked gZiPro exhibits clear chemical shift differences for many residues when the ^{1}H-^{15}N-HSQC spectra of bZiPro and gZiPro are compared upon binding to a dipeptide inhibitor [35]. Another study suggests that exchanges might be present in the linked protease (gZiPro) because both open and closed conformations can be observed in solution [24]. Taken together, all these protease constructs are well folded in solution despite different dissociation events observed between NS2B and NS3 in the bZiPro and gZiPro constructs. Inhibitors or substrates can further stabilize protease conformations in solution. The residues from the N-terminal region of NS3 are flexible and detached from the active site after cleavage. Presence of the glycine linker affects the chemical environments of residues, which might affect inhibitor binding and sample homogeneity [35]. The unlinked protease (bZiPro) is suitable for screening inhibitors with different molecular weights and potencies.

Finding ZIKV NS2B-NS3 Protease Inhibitors

Developing inhibitors against proteases of flaviviruses has been of great interest for researchers from both academia and industry. Several potent inhibitors against West Nile and dengue proteases have been developed, but none of them has entered into clinical trials because they lack cellular activity and exhibit poor pharmacokinetics [14, 36, 37]. The design of protease inhibitors with good oral bioavailability is often difficult and the flat and highly charged active site of flavivirus NS2B-NS3 proteases increases

the challenges of inhibitor design [38]. These difficulties can be mitigated by structural studies as has been demonstrated by the development of HCV protease inhibitors [39].

The best starting point for the drug discovery of serine proteases are usually the substrate peptides [38]. Such peptides can be turned into nanomolar inhibitors through the addition of electrophilic warheads like boronic acids [30] or β-keto amides [40]. While this technique provides very potent inhibitors, it has no effect on the key challenges of peptidic inhibitors, namely cellular activity and oral pharmacokinetics. The only available strategy to turn polar peptides into potential drugs is the stepwise transformation into peptidomimetics which have reduced molecular weight and improved properties. Such optimizations benefit greatly from co-crystal structures with inhibitors, in fact it is not unusual that during such a project 20-30 enzyme inhibitor structures are used to facilitate computational drug design.

As mentioned previously, this strategy could not be employed for the dengue or West Nile NS2B-NS3 proteases since X-ray crystal structures of the closed conformation without an inhibitor could not be obtained. Furthermore it was never possible to crystallize multiple structures of related peptides to study the effect of reducing the peptide content for the potent peptide inhibitors. This is very different for the Zika NS2B-NS3 protease. After a relatively short time the research community has produced multiple structures of different forms of the protease and very importantly protease structures without inhibitors have become available [22, 25, 26, 29, 30, 32]. As a consequence, the availability of reproducible, strongly diffracting and soakable crystals augurs well for Zika virus drug discovery [26]. Nevertheless the literature on DENV protease inhibitors suggests that the design of cell permeable ZIKV protease inhibitors will be challenging [14, 41-43].

The search for small molecular weight compounds of flaviviral NS2B-NS3 protease inhibitors has been fraught with difficulties. While many authors have reported the identification of weak inhibitors from high throughput screening (HTS), further optimization has invariably been unsuccessful. In a recent report, HTS against a library containing more than 40, 000 compounds identified some inhibitors of the HCV NS3/NS4A protease. Ten of the identified inhibitors exhibited activities against the ZIKV

protease. Further biophysical and structural studies confirmed the binding for the available compounds, which provides a structural basis for compound optimization [29]. Another study identified bromocriptine to be an inhibitor of ZIKV protease in vitro and the authors proposed that combination of bromocriptine and interferon could be a potential treatment for ZIKV infection [44].

Fragment-based drug discovery (FBDD) has only rarely been used for protease drug discovery because of the challenges of obtaining information about the active site interactions of such weak binding compounds. Surprisingly a fragment hit (EN300) was identified and was confirmed to bind to the protease active site by crystallography [26]. This observation is encouraging because it again suggests that the Zika NS2B-NS3 protease is well suited for studying the molecular recognition of inhibitors.

The major hurdle for cellular activity is membrane permeation, which is mainly impeded by high molecular weight and high charge/polarity of the ZIKV inhibitor. This is a standard problem for protease inhibitors that mimic substrate peptides. In order to solve this problem for example CaCO2 cell permeability would be a useful preclinical assay. In addition any ZIKV cellular assay (qRT-PCR viral yield or plaque assay) would be useful to gage the cellular activity of the compounds. Brecher et al. recently reported a conformational switch assay based on the split luciferase complementation to monitor the conformational change of NS2B and to characterize candidate allosteric inhibitors, which represents a convenient cell based assay for compound screening [45].

Concluding Remarks

HTS campaigns, virtual screening and peptidomimetics have been used to design dengue protease inhibitors. That no inhibitor has reached the clinical stage so far for dengue is probably due to the low druggability of the active site. Although the Zika protease is structurally similar to the dengue protease, accumulated structural studies suggest that there is a greater likelihood that Zika protease inhibitors can be developed. Recent progress in biophysical studies on Zika protease highlights its key advantages over the dengue and West Nile proteases for designing antivirals. First, detailed

structural and biochemical studies have been carried out for the three distinct protease constructs, which answers the questions raised from the West Nile and dengue protease structural studies [13, 14]. The open conformation is not suitable for structure-based drug design as it is only observed for the linked proteases. Second, structures of the protease-substrate complexes are available, which was not achieved for either the West Nile or dengue proteases. Third, FBDD is promising for the Zika protease as a fragment hit was identified and verified by biophysical and structural studies. Such progress was not made for dengue and West Nile proteases. Fourth, the unlinked protease is likely to be an ideal construct for structure-based drug design because it is been demonstrated to exist in the closed conformation and possess an open active site accessible to different types of inhibitors ranging from high molecular weight (BPTI) inhibitors to low molecular weight fragments. Last, eZiPro structure represents a native conformation of the protease-product complex. This construct might have more useful applications such as screening fragment hits targeting different sites. Structure-based drug discovery can be a very practical strategy to develop inhibitors, and has great potential for its application in developing inhibitors of ZIKV (see Outstanding Questions).

Acknowledgements

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.

References

1 Ferguson, N.M.*, et al.* (2016) EPIDEMIOLOGY. Countering the Zika epidemic in Latin America. *Science* 353, 353-354

2 Heymann, D.L.*, et al.* (2016) Zika virus and microcephaly: why is this situation a PHEIC? *Lancet* 387, 719-721

3 Wilder-Smith, A.*, et al.* (2017) Epidemic arboviral diseases: priorities for research and public health. *Lancet Infect Dis* 17, e101-e106

4 Russell, P.K. (2016) The Zika Pandemic - A Perfect Storm? *PLoS Negl Trop Dis* 10, e0004589

5 de Araujo, T.V.*, et al.* (2016) Association between Zika virus infection and microcephaly in Brazil, January to May, 2016: preliminary report of a case-control study. *Lancet Infect Dis* 16, 1356-1363

6 Bardina, S.V.*, et al.* (2017) Enhancement of Zika virus pathogenesis by preexisting antiflavivirus immunity. *Science* 356, 175-180

7 Malone, R.W.*, et al.* (2016) Zika Virus: Medical Countermeasure Development Challenges. *PLoS Negl Trop Dis* 10, e0004530

8 Fernandez, E. and Diamond, M.S. (2017) Vaccination strategies against Zika virus. *Current opinion in virology* 23, 59-67

9 Barrows, N.J.*, et al.* (2016) A Screen of FDA-Approved Drugs for Inhibitors of Zika Virus Infection. *Cell host & microbe* 20, 259-270

10 Xu, M.*, et al.* (2016) Identification of small-molecule inhibitors of Zika virus infection and induced neural cell death via a drug repurposing screen. *Nat Med* 22, 1101-1107

11 Huang, Q.*, et al.* (2013) Lyso-myristoyl phosphatidylcholine micelles sustain the activity of Dengue non-structural (NS) protein 3 protease domain fused with the full-length NS2B. *Protein expression and purification* 92, 156-162

12 Liew, L.S.*, et al.* (2016) Selection of suitable detergents for obtaining an active dengue protease in its natural form from *E. coli. Protein Expr Purif* 121, 141-148

13 Luo, D.*, et al.* (2015) The flavivirus NS2B-NS3 protease-helicase as a target for antiviral drug development. *Antiviral Res* 118, 148-158

14 Nitsche, C.*, et al.* (2014) Biochemistry and medicinal chemistry of the dengue virus protease. *Chemical reviews* 114, 11348-11381

15 Noble, C.G.*, et al.* (2012) Ligand-bound structures of the dengue virus protease reveal the active conformation. *J Virol* 86, 438-446

16 Robin, G.*, et al.* (2009) Structure of West Nile virus NS3 protease: ligand stabilization of the catalytic conformation. *J Mol Biol* 385, 1568-1577

17 Luo, D.*, et al.* (2008) Crystal structure of the NS3 protease-helicase from dengue virus. *J Virol* 82, 173-183

18 Erbel, P.*, et al.* (2006) Structural basis for the activation of flaviviral NS3 proteases from dengue and West Nile virus. *Nat Struct Mol Biol* 13, 372-373

19 Kim, Y.M.*, et al.* (2013) NMR analysis of a novel enzymatically active unlinked dengue NS2B-NS3 protease complex. *J Biol Chem* 288, 12891-12900

20 Gruba, N.*, et al.* (2016) Substrate profiling of Zika virus NS2B-NS3 protease. *FEBS letters* 590, 3459-3468

21 Rut, W.*, et al.* (2017) Extended substrate specificity and first potent irreversible inhibitor/activity-based probe design for Zika virus NS2B-NS3 protease. *Antiviral Res* 139, 88-94

22 Phoo, W.W.*, et al.* (2016) Structure of the NS2B-NS3 protease from Zika virus after self-cleavage. *Nature communications* 7, 13410

23 Kuiper, B.D.*, et al.* (2017) Increased activity of unlinked Zika virus NS2B/NS3 protease compared to linked Zika virus protease. *Biochem Biophys Res Commun* doi: 10.1016/j.bbrc.2017.03.108

24 Mahawaththa, M.C., *et al.* (2017) Solution conformations of a linked construct of the Zika virus NS2B-NS3 protease. *Antiviral Res* 142, 141-147
25 Li, Y., *et al.* (2017) Structural dynamics of Zika virus NS2B-NS3 protease binding to dipeptide inhibitors. *Structure* 25, 1-9
26 Zhang, Z., *et al.* (2016) Crystal structure of unlinked NS2B-NS3 protease from Zika virus. *Science* 354, 1597-1600
27 Choksupmanee, O., *et al.* (2012) Structural platform for the autolytic activity of an intact NS2B-NS3 protease complex from dengue virus. *Biochemistry* 51, 2840-2851
28 Li, Y., *et al.* (2015) Membrane topology of NS2B of dengue virus revealed by NMR spectroscopy. *Biochim Biophys Acta* 1848, 2244-2252
29 Lee, H., *et al.* (2017) Identification of novel small molecule inhibitors against NS2B/NS3 serine protease from Zika virus. *Antiviral Res* 139, 49-58
30 Lei, J., *et al.* (2016) Crystal structure of Zika virus NS2B-NS3 protease in complex with a boronate inhibitor. *Science* 353, 503-505
31 Leon-Juarez, M., *et al.* (2016) Recombinant Dengue virus protein NS2B alters membrane permeability in different membrane models. *Virology journal* 13, 1
32 Chen, X., *et al.* (2016) Mechanisms of activation and inhibition of Zika virus NS2B-NS3 protease. *Cell research* 26, 1260-1263
33 Su, X.C., *et al.* (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 neglected tropical diseases* 3, e561
34 de la Cruz, L., *et al.* (2011) Binding of low molecular weight inhibitors promotes large conformational changes in the dengue virus NS2B-NS3 protease: fold analysis by pseudocontact shifts. *J Am Chem Soc* 133, 19205-19215
35 Li, Y., *et al.* (2017) Structural characterization of the linked NS2B-NS3 protease of Zika virus. *FEBS letters*
36 Yin, Z., *et al.* (2006) Peptide inhibitors of Dengue virus NS3 protease. Part 1: Warhead. *Bioorganic & medicinal chemistry letters* 16, 36-39
37 Yin, Z., *et al.* (2006) Peptide inhibitors of dengue virus NS3 protease. Part 2: SAR study of tetrapeptide aldehyde inhibitors. *Bioorganic & medicinal chemistry letters* 16, 40-43
38 Poulsen, A., *et al.* (2014) Drug design for flavivirus proteases: what are we missing? *Curr Pharm Des* 20, 3422-3427
39 Njoroge, F.G., *et al.* (2008) Challenges in modern drug discovery: a case study of boceprevir, an HCV protease inhibitor for the treatment of hepatitis C virus infection. *Accounts of chemical research* 41, 50-59
40 Steuer, C., *et al.* (2011) Synthesis and biological evaluation of alpha-ketoamides as inhibitors of the Dengue virus protease with antiviral activity in cell-culture. *Bioorg Med Chem* 19, 4067-4074
41 Weigel, L.F., *et al.* (2015) Phenylalanine and Phenylglycine Analogues as Arginine Mimetics in Dengue Protease Inhibitors. *Journal of medicinal chemistry* 58, 7719-7733
42 Behnam, M.A., *et al.* (2015) Discovery of Nanomolar Dengue and West Nile Virus Protease Inhibitors Containing a 4-Benzyloxyphenylglycine Residue. *Journal of medicinal chemistry* 58, 9354-9370
43 Kouretova, J., *et al.* (2017) Effects of NS2B-NS3 protease and furin inhibition on West Nile and Dengue virus replication. *Journal of enzyme inhibition and medicinal chemistry* 32, 712-721
44 Chan, J.F., *et al.* (2017) Novel antiviral activity and mechanism of bromocriptine as a Zika virus NS2B-NS3 protease inhibitor. *Antiviral Res* 141, 29-37
45 Brecher, M., *et al.* (2017) A conformational switch high-throughput screening assay and allosteric inhibition of the flavivirus NS2B-NS3 protease. *PLoS pathogens* 13, e1006411

46 Lee, H.*, et al.* (2016) Identification of novel small molecule inhibitors against NS2B/NS3 serine protease from Zika virus. *Antiviral research* 139, 49-58
47 Kang, C.*, et al.* (2013) Exploring the binding of peptidic West Nile virus NS2B-NS3 protease inhibitors by NMR. *Antiviral Res* 97, 137-144

Glossary

- **Serine Proteases** (or serine endopeptidases) are enzymes that cleave peptide bonds in proteins, which serine serves as the nucleophilic amino acid at the (enzyme's) active site. They are found ubiquitously in both eukaryotes and prokaryotes. For flaviviral proteases, Ser135 from NS3 protein is the catalytic serine.
- **Peptidomimetic inhibitors** are derived from the protease cleavage sites. This type of protease inhibitors contains basic residues which facilitate its interactions with NS2B-NS3 protease. To enhance the inhibitor activity, additional reactive group and chemical modifications are usually required.
- **Small molecular weight inhibitors** are low molecular weight (< 900 daltons) organic compounds that inhibit the protease activity.
- **High-throughput screening (HTS)** is a method used in drug discovery. Using robotics, data processing/control software, liquid handling devices, and sensitive detectors, high-throughput screening allows a researcher to quickly conduct millions of chemical, genetic, or pharmacological tests in order to rapidly identify active compounds. The lead compounds serve as starting points for drug design and for understanding the interaction or role of a particular biochemical process in biology.
- **Fragment-based drug design (FBDD)** is a method used for finding lead compounds as part of the drug discovery process. It is based on identifying small chemical fragments, which may bind only weakly to the biological target, and then growing them or combining them to produce a lead with a higher affinity. FBDD can be compared with high-throughput screening (HTS). In HTS, libraries with up to millions of compounds, with molecular weights of around 500 Da, are screened, and nanomolar binding affinities are sought. In contrast, in the early phase of FBDD, libraries with a few thousand compounds with molecular weights of around 200 Da may be screened, and millimolar affinities can be considered useful.

Box 1. Protein Constructs of ZIKV NS2B-NS3 Protease

A single-chain construct where the NS2B cofactor region is linked to NS3 protease region through a glycine-rich ($Gly_4SerGly_4$) linker has been used for earlier flaviviral protease studies. For all reported structures using this construct design, the free enzyme has been captured in an open conformation where the NS2B catalytic cofactor region is disordered and the active site is not completely formed. Enzyme-inhibitor complexes form the closed conformation where the NS2B cofactor is fully engaged with NS3 protease for inhibitor recognition at P2 position.

When structural work on Zika virus NS2B-NS3 protease started, three different protein and protein complexes were examined (**Fig. 1**) [22]. The first one is using the traditional approach to connect the central cofactor region of NS2B to the NS3 protease domain via the nine-residue Gly4SerGly4 linker, providing a single-chain protein which was named as **gZiPro**. The second approach highlights the enzymatic self-cleavage of the NS2B-NS3 junction. This "**eZiPro**" construct is cleaved during protein expression providing an unlinked protein complex [22]. Finally "**bZiPro**" stands for a bivalent protease consisting of two separate polypeptide fragments 45-96 of NS2B and 1-177 of NS3 co-expressed and assembled in bacteria [22]. All three proteases fold as monomeric globular proteins and are catalytically active in solution. However, owing to the differences of the constructs, the protease activity and kinetic parameters of the three proteases are different. Based on the available data, bZiPro design seems to offer several advantages for structure based inhibitor development: (1) The free enzyme maintains the closed conformation in solution and in crystals; (2) The active site is preformed and fully accessible to inhibitors of different sizes and binding affinities; (3) High resolution crystal structures of free enzyme and compound-bound complexes are available. (4) Most protein residues are well resolved in both crystal structures and in 2D-NMR assignment.

Table 1. ZIKV NS2B-NS3 protease structures [a]				
Design	**PDB code**	**Ligand**	**Resolution**	**Reference**
eZiPro	5GJ4	TGKR peptide	1.84 Å	[22]
gZiPro	5LC0	Boronate dipeptide inhibitor	2.7 Å	[30]
	5T1V	-	3.1Å	[46]
	5GXJ	-	2.6 Å	[32]
bZiPro	5GPI	-	1.58 Å	[26]
		KKGE reverse peptide		
	5H4I	benzimidazol-1-ylmethanol	2.0 Å	[26]
	5H6V	Acyl-KR-aldehyde	2.42 Å	[25]

[a] Structures of proteases from other flaviviruses have been summarized earlier [13, 14].

Figure Legends

Figure 1. ZIKV polyprotein and NS2B-NS3 protease. (**A**) The ten viral proteins are shown in different colors and the cleavage sites for the NS2B-NS3 protease and host proteases are marked with solid arrows and hollow arrows, respectively. (**B**) The protease cleavage sequences are shown for the three constructs of the NS2B-NS3 protease. The representative crystal structures of (**C**) bZiPro free enzyme, (**D**) eZiPro, (**E**) gZiPro-inhibitor complex, and (**F**) gZiPro free enzyme. Color scheme follows (**B**). The missing inter-domain residues are shown as dotted lines. One dot corresponds to one missing residues in the crystal structures. (**G**) Membrane associated model of ZIKV NS2B-NS3 Protease. The transmembrane region of NS2B was based on the NMR work on DENV NS2B [28].

Figure 2. The catalytic site of the ZIKV NS2B-NS3 protease. (**A**) Essential interactions between the protease and a dipeptide KR-COOH, using the eZiPro structure as reference. Key interactions are labelled and NS2B residue labels are underlined. The negatively charged residue D83 from NS2B from the β-hairpin interacts with the side chain of P2 residue (K129) in the substrate. Residue R130 in the substrate occupies the S1 pocket and forms interactions with Y161, D129 and Y130 from NS3. The P3 and P4 residues in the substrate play little role in binding as the S3 and S4 pockets are empty, which explains the recent finding that the dipeptide inhibitor can exhibit an IC50 in the nanomolar range against the WNV protease[47]. NS3 protein is colored in yellow and NS2B in magenta. Hydrogen bonds are shown as dashes. Surface charge density view of the protease-inhibitor complexes for (**B**) eZiPro (TGKR-COOH peptide), (**C**) Boronate dipeptide inhibitor, (**D**) KKGE reverse peptide, (**E**) Acyl-KR-aldehyde, and (**F**) small fragment benzimidazol-1-ylmethanol. Positively charged surfaces are in blue and negatively charged in red. S1-S4 substrate binding pockets are labelled in B. (**G**) Model of NS2B-NS3 on a membrane. ZIKV NS2B membrane topology was modelled based on DENV NS2B [28].

Figure 3. NMR studies on Zika protease. (**A**) $_{1}$H-$_{15}$N-HSQC spectra of Zika protease. The $_{1}$H-$_{15}$N-HSQC spectra of bZiPro (red), eZiPro (black), and gZiPro (green) are shown. The spectra were collected at 37 °C [22]. (**B**) Residues that have exchanges in

the Zika protease. The crystal structure of bZiPro in complex with Ac-KR-aldehyde (Ac-KR-COH; PDB id 5H6V) is shown. Ac-KR-aldehyde is shown in sticks. NS2B and NS3 are shown in purple and yellow, respectively. Residues exhibiting line broadening in free bZiPro are highlighted in light blue. (**C**) bZiPro binds to BPTI [22]. (**D**) bZiPro binds to a weak peptide inhibitor Ac-KR-COOH [26]. ^{1}H-^{15}N-HSQC spectra of bZiPro in the absence (black) and presence (red) of Ac-KR-COOH are shown.

Figure 1

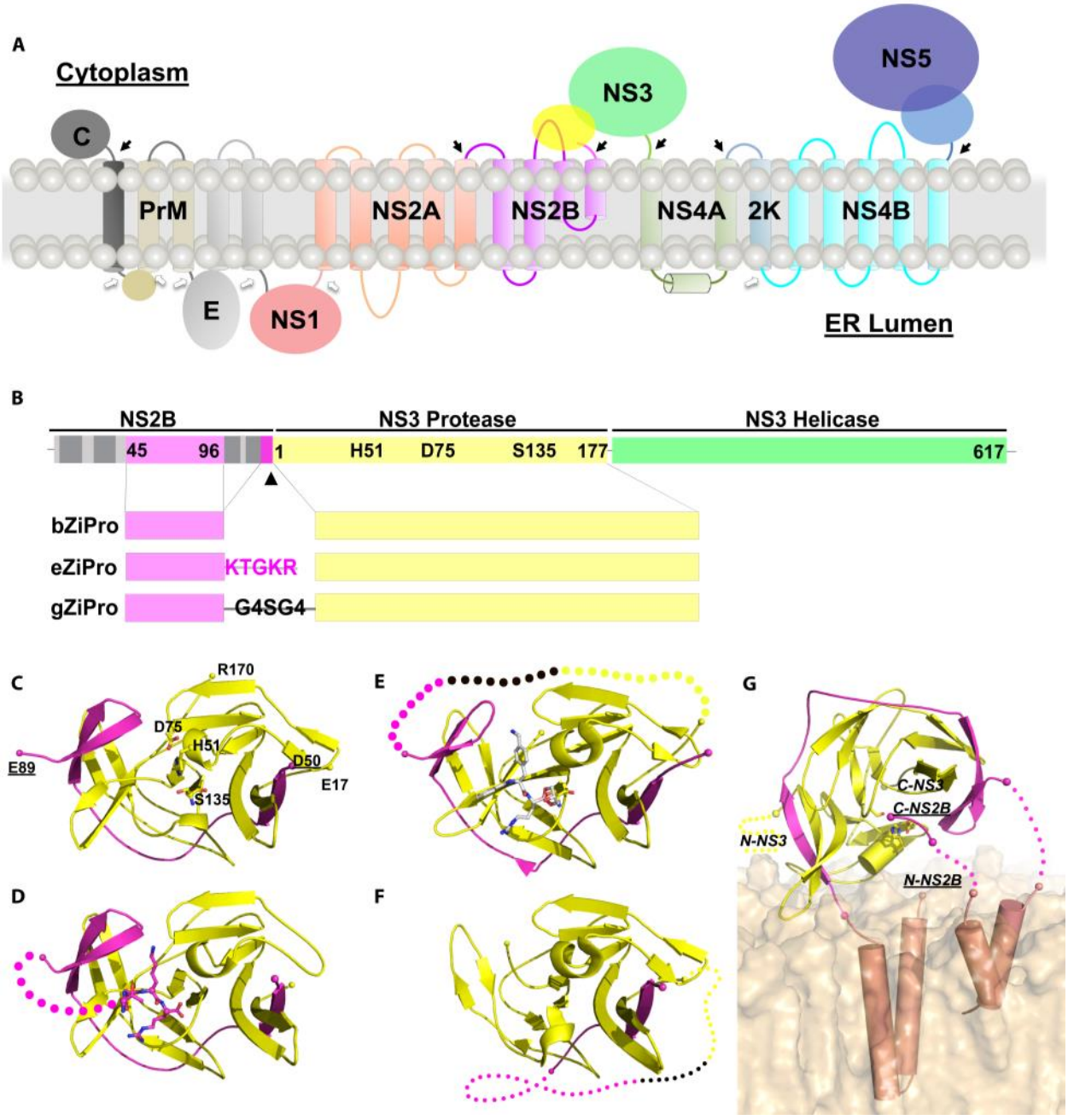

A
Cytoplasm
C
PrM
E
NS1
NS2A
NS2B
NS3
NS4A
2K
NS5
NS4B
ER Lumen
B
NS2B
NS3 Protease
NS3 Helicase
45
96
1
H51
D75
S135
177
617
bZiPro
eZiPro
KTGKR
gZiPro
G4SG4
C
R170
D75
H51
E89
D50
E17
S135
D
E
F
G
C-NS3
C-NS2B
N-NS3
N-NS2B

Figure 2

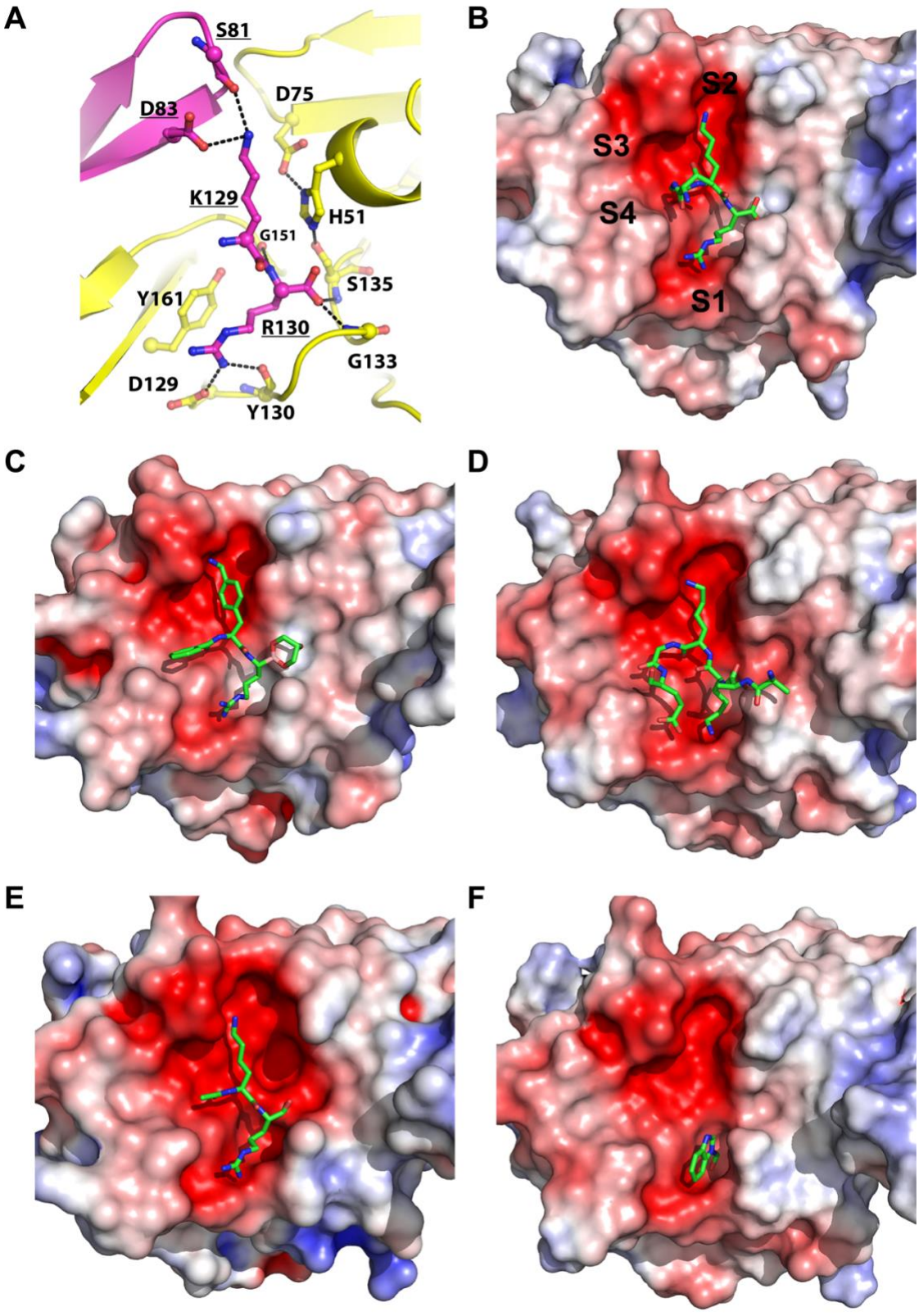
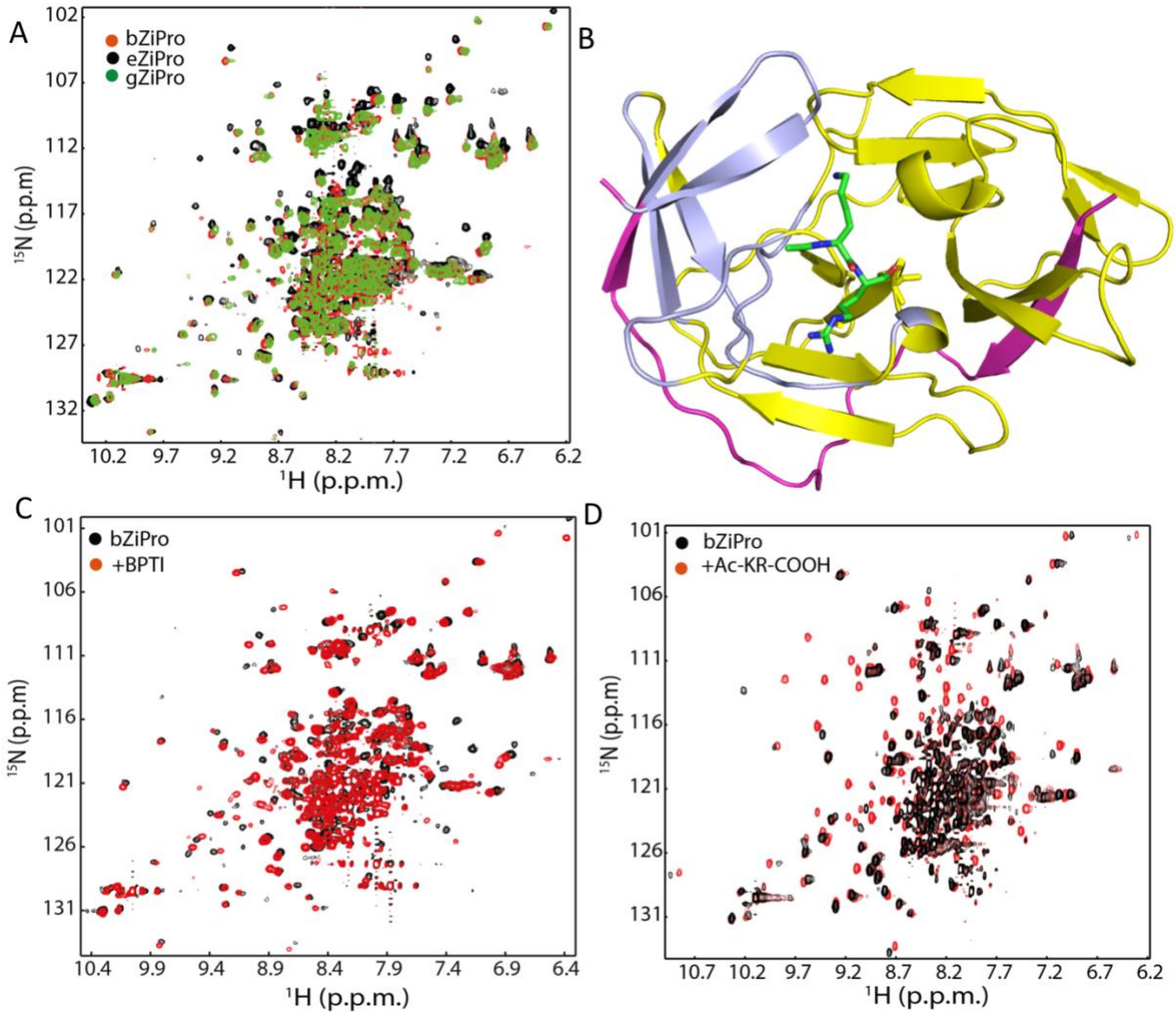

A
S81
D83
D75
K129
H51
G151
S135
Y161
R130
G133
D129
Y130
B
S2
S3
S4
S1
C
D
E
F

Figure 3


A
bZiPro
eZiPro
gZiPro
^{15}N (p.p.m)
^{1}H (p.p.m.)
B
C
bZiPro
+BPTI
^{15}N (p.p.m)
^{1}H (p.p.m.)
D
bZiPro
+Ac-KR-COOH
^{15}N (p.p.m)
^{1}H (p.p.m.)