frontiers | Frontiers in Medicine

ORIGINAL RESEARCH
published: 06 July 2022
doi: 10.3389/fmed.2022.921060

C-Terminal Extended Hexapeptides as Potent Inhibitors of the NS2B-NS3 Protease of the ZIKA Virus

Suyash Pant[1] *and Nihar R. Jena*[2*]

[1] Department of Pharmacoinformatics, National Institute of Pharmaceutical Education and Research, Kolkata, India,
[2] Discipline of Natural Sciences, Indian Institute of Information Technology, Design and Manufacturing, Jabalpur, India

OPEN ACCESS

Edited by:
Hu Wang,
Johns Hopkins University,
United States

Reviewed by:
Tuoxian Tang,
University of Pennsylvania,
United States
Yang Chen,
Chinese Academy of Sciences
(CAS), China
Ye Wang,
Georgia State University,
United States
Yi Wang,
Sichuan Academy of Medical
Sciences and Sichuan Provincial
People's Hospital, China
Chuanqi Sun,
UCLA Health System, United States
Hao Yan,
Stanford University, United States
Ruipu Xin,
UCLA Health System, United States

***Correspondence:**
Nihar R. Jena
nrjena@iiitdmj.ac.in

Specialty section:
This article was submitted to Translational Medicine, a section of the journal Frontiers in Medicine

Received: *15 April 2022*
Accepted: *07 June 2022*
Published: *06 July 2022*

Citation:
Pant S and Jena NR (2022) C-Terminal Extended Hexapeptides as Potent Inhibitors of the NS2B-NS3 Protease of the ZIKA Virus. Front. Med. 9:921060. doi: 10.3389/fmed.2022.921060

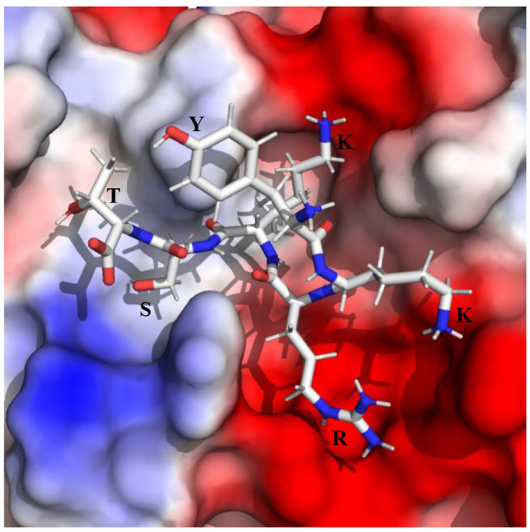
The Zika virus (ZIKV) protease is an attractive drug target for the design of novel inhibitors to control the ZIKV infection. As the protease substrate-binding site contains acidic residues, inhibitors with basic residues can be beneficial for the inhibition of protease activities. Molecular dynamics (MD) simulation and molecular mechanics with generalized Born and surface area solvation (MM/GBSA) techniques are employed herein to design potent peptide inhibitors and to understand the nature of the basic residues that can potentially stabilize the acidic residues of the protease substrate-binding site. It is found that the inclusion of K, R, and K at P1, P2, and P3 positions, respectively, and Y at the P4 position (YKRK) would generate a highly stable tetrapeptide-protease complex with a ΔG_{bind} of ~ -80 kcal/mol. We have also shown that the C-terminal extension of this and the second most stable tetrapeptide (YRRR) with small polar residues, such as S and T would generate even more stable hexapeptide-protease complexes. The modes of interactions of these inhibitors are discussed in detail, which are in agreement with earlier experimental studies. Thus, this study is expected to aid in the design of novel antiviral drugs against the ZIKV.

Keywords: Zika virus, NS2B–NS3 protease, peptide inhibitors, covalent inhibitors, MD-simulations, peptidomimetics

INTRODUCTION

The Zika virus (ZIKV) infection causes both mild and severe diseases including fever, joint pain, (1, 2), Guillain-Barre syndrome (3), acute myelitis (4), and brainstem dysfunctions (5). In the case of infected pregnant women, it induces microcephaly (6), and other congenital malformations (7–9). Although ZIKV infection was declared a global emergency by the World health organization (WHO) in the year 2016 (10), no approved vaccine or drug is available to date to contain this disease (11–13).

The ZIKV contains a single-stranded RNA genome that translates to form a polypeptide chain inside host cells. Subsequently, this polypeptide gets cleaved by the host and viral proteases to form three structural [Envelop (E), membrane (M), and capsid (C)] and seven non-structural (NS1, NS2A, NS2B, NS3, NS4A, NS4B, NS5) proteins (14). Among these proteins, the NS3 encodes the serine protease, RNA helicase, RNA 5' triphosphatase (RTPase), and nucleocapsid triphosphatase (NTPase) activities (14–17). The N-terminal region of NS3 in association with the membrane-bound NS2B co-factor constitutes the serine protease (14, 15), which is responsible for the cleavage of the viral polyprotein and different key host proteins involved in immune response



GRAPHICAL ABSTRACT | Binding of YKRKST to the protease, which fully occupies its substrate-binding site and hence may act as a potent inhibitor of the ZIKA virus protease.

(18). Therefore, it is necessary to design substrate competitive inhibitors to inhibit the protease activities of NS2B–NS3 protease to contain the ZIKV (19–35).

Several attempts were made to design substrate competitive inhibitors to occupy the substrate-binding site of the NS2B–NS3 protease of the ZIKV. These inhibitors include (1) small molecules (19–27), (2) peptidomimetics (28–33), and (3) peptide inhibitors (34, 35). As the protease active site is surrounded by negatively charged amino acids (acidic residues), the binding of small molecule neutral inhibitors was not effective. However, peptidomimetic inhibitors that contain positively charged amino acids (basic residues), such as Lys (K) and Arg (R), and different organic warheads were proposed to be efficient (29, 30). These inhibitors can either make a covalent bond with the catalytically important residue Ser135 of NS3 (28, 29) or bind non-covalently to the protease (30). Interestingly, recently, non-covalent peptide inhibitors were shown to bind strongly with the NS2B–NS3 protease (35). Among these inhibitors, the YKKR was found to possess the highest binding free energy (35). It was shown that the P1 R binds to the S1 substrate site, while P2 K and P3 K bind to the S2 and S3 sites, respectively. It was proposed that the heavy amino acid Y at the P4 position provides the necessary conformational rigidity to KKR to fully occupy the substrate-binding site (35). However, for the Dengue virus (DENV) protease, strong preferences for both K and R were reported at the P1 position, whereas, R and K were preferred at the P2 and P3 positions, respectively (36). For the West Nile Virus (WNV), K/R at the P1 position and K at both P2 and P3 positions were found to be preferred (37, 38). As the DNV and WNV proteases are structurally similar to that of the ZIKV (39), it is necessary to understand the binding preferences of K and R at different substrate sites of the ZIKV protease. This will eventually help in the development of potent inhibitors for the inhibition of ZIKV protease activities.

Although the bindings of different inhibitors to the unprimed sites (S1–S4) are well studied (19–35), the inhibitor binding to the prime sites (S1' and S2') is rarely studied. Therefore, it is also desirable to understand the effects of prime site residues on the stability of peptide-protease complexes. Analysis of sequences of different flavivirus substrates indicates that Ser (S) and Thr (T) are present at the P1' and P2' positions respectively (36). As these polar residues are small in size, they can be well accommodated in the small cleft made by the S1' and S2' substrate sites of the protease. Therefore, the C-terminal extension of the most stable tetrapeptide inhibitors by S and T may further stabilize the peptide-protease complexes.

To identify a promiscuous inhibitor of the ZIKV protease, the positions of K and R in the YKKR-protease complex were changed without perturbing the P4 Y to generate 7 different peptides. Subsequently, the structural and dynamical effects of these inhibitors bound to the protease were undertaken to elucidate their roles in creating closed complex structures. Eventually, the relative Gibbs binding free energy analysis was carried out to short-list the most stable peptide-protease complex. The two most stable complexes were further extended by adding S and T to their C-terminal ends to understand the effects of prime residues on the inhibitor binding.

COMPUTATIONAL METHODOLOGY

System Preparation

Recently, the binding of YKKR to the bZipro form of the protease (40) was shown to produce a stable complex with a relative binding free energy of about -73 ± 8 kcal/mol (35). The average simulated structure of the YKKR-protease complex was found to be similar to the X-ray structures of Acyl-KR-Aldehyde-protease (PDB ID 5H6V) (29), Phenylacetyl-KKR-protease (PDB ID 5ZMQ) (30), and TGKR-protease (PDB ID 5GJ4) (34) complexes. For this reason, the average simulated complex structure of the YKKR (35) was converted to seven different peptides, such as YKKK, YKRK, YRKK, YKRR, YRKR, YRRK, and YRRR by mutating P1 R, P2 K, and P3 K without perturbing the initial backbone conformation of YKKR. The PyMOL program (41) was used to create the mutated peptides. As these peptides were subjected to molecular dynamics simulations, it is believed that the mutated structures are not biased to their initial conformation (35, 42, 43).

Molecular Dynamics (MD) Simulations and MM/GBSA Calculations

The Desmond 2021-1 program of Schrodinger (44, 45) was used to solvate the peptide-protein complexes by placing them in an explicit water box of size 10 Å. The OPLS4 force field (46) was used to model the peptide inhibitors. The partial charges of the ligands were generated by using the same force field. The single-point charge (SPC) model (47) was used to account for the explicit water molecules. Sufficient numbers of ions were added to make the solvated complexes neutral. The protonation states of the protein and peptidomimetic ligands were set as

per the $pH = 7.4$. Subsequently, these complexes were energy minimized by 2,000 steps each by using the steepest descent and limited-memory Broyden-Fletcher-Goldfarb-Shanno (LBFGS) algorithms (48). The minimized complexes were slowly heated to maintain a temperature of 300 kelvin (K) in several steps by using the Nose-Hoover thermostatic algorithm (49). In the first step, the system was heated to 10 K for 100 ps. to reduce any possible steric clashes. In the second step, a 12 ps. of molecular dynamics run was performed with the NVT ensemble to relax the system at 10 K. In the third step, molecular dynamics run of 12 ps. was carried out by using the NPT ensemble, where a pressure of 1 atm was maintained by using the Langevin barostat (50). In the fourth step, the temperature was raised to 300 K for 12 ps. by using the NPT ensemble. In all of the above steps, the solute heavy atoms were restrained with a force constant of 50 kcal $mol^{-1}Å^{-1}$. In the fifth step, restraint was released and the molecular dynamics simulation was carried out at the NPT ensemble for 24 ps. Consequently, all complexes were subjected to a production run for 200 ns by considering the integration time step of 1 fs and the NPT ensemble. The periodic boundary condition (PBC) (51) was considered for all of the simulations.

To calculate the relative binding free energy (ΔG_{bind}) of each protease-inhibitor complex, the MM/GBSA technique as implemented in the Desmond 2021-1 package (45, 46) was used. For this purpose, 100 snapshots were extracted from the last 10 ns trajectories of each complex at an interval of 100 ps. Equation (1) was used to compute ΔG_{bind}.

$$G_{bind} = G_{complex\ (minimized)} - G_{protein\ (unbound,minimized)} - G_{ligand\ (unbound,minimized)} \quad (1)$$

where, $G_{complex(minimized)}$ is the MM/GBSA energy of the minimized complex, $G_{protein(unbound,minimized)}$ is the MM/GBSA energy of the minimized protein after separating it from its bound ligand and $G_{lig(unbound,minimized)}$ is the MM/GBSA energy of the ligand after separating it from the complex and allowing it to relax. However, as entropy calculations were not performed, the free energy terms contain contributions from enthalpy only.

RESULTS AND DISCUSSIONS

The root mean square deviations (RMSD) of the Cα atoms of the protease and the root mean square fluctuations of different residues of the protease evolved during the simulations are illustrated in **Figure 1**. As the average RMSD of each complex is < 2.5 Å, the complexes were stable during the simulations. The interaction diagrams (including hydrogen bonding, electrostatic, π-cation, π-π interactions, etc.) depicted in **Supplementary Figures S1–S10** elucidates the detailed interactions of the peptide inhibitors with the protease evolved during the simulations. The interactions that lasted for ≥80, 50–79, and <50% of the simulation time are considered to be strong, moderate, and weak, respectively (49). It should be mentioned that a hydrogen bonding interaction satisfies the following geometric criteria : (1) The protein-ligand H-bond distance is ≤2.5 Å between the donor and acceptor atoms (D–H···A), (2) a donor angle of ≥120° exists between the donor-hydrogen-acceptor atoms (D–H···A), and (3) an acceptor angle of ≥90° exists between the hydrogen-acceptor-bonded atoms (H···A–X). The π-cation interaction is defined as the interaction between an aromatic and charged groups situated within a distance of 4.5 Å. Similarly, the π-π interaction is defined as the interaction between two aromatic groups stacked face-to-face or face-to-edge by maintaining a distance of <4.5 Å. The relative binding free energies (ΔG_{bind}) of different peptide-protease complexes presented in **Table 1** indicate the binding strength of different peptide inhibitors studied herein.

The Most Stable Peptide-Protease Complex

Most of the peptidomimetics that contain di- or tripeptides have R at the P1 position and K at the P2 position (28–31). We have also recently shown that YKKR tetrapeptide with R and K at the P1 and P2 positions, respectively binds strongly to the protease (35). However, the consideration of alternate combinations of R and K at P1, P2, and P3 positions revealed that the stabilities of different complexes involving tetrapeptides would follow the order YKRK > YRRR > YRRK >YKK ≥ YKKR > YKRR >YKKK > YRKR (**Table 1**). This indicates that the YKRK-protease complex would be the most stable one, which is about 8 kcal/mol more stable than the YKKR-protease complex (**Table 1**). The YKRK peptide contains P1 K instead of P1 R, clamped by P2 R and P3 K. It is also about 3 kcal/mol more stable than the second most stable complex of this series (YRRR-protease). As both YKRK and YRRR contain P2 R and P4 Y and the P3 residue does not contribute to the acid-base interaction, the slightly higher stability of the former complex is likely due to the stronger binding of P1 K with the residues of S1 site. This is evident from **Figure 2** and **Supplementary Figure S1**, where P1 K is more suitably placed in the S1 site to make a salt-bridge interaction with Asp129 (64% occupancy), a hydrogen bond with Tyr130 (68% occupancy), and a π-cation interaction with Tyr160 (81% occupancy) of NS3 (**Figure 2A**). Additionally, its backbone amide makes a strong hydrogen bond with Gly151 (92% occupancy) of NS3. However, in the case of YRRR, the positively charged guanidine group of P1 R points away from Asp129 and missed the key ionic interaction with Asp129 (**Figure 2B**). Further, the C-terminal carboxyl group of YKRK makes direct and indirect hydrogen bonds with S135, Gly133, and Val36 (<50% occupancy) of NS3 (**Figure 2A**), which are missing in YRRR (**Figure 2B**). Similarly, the side chain of P2 R in YKRK makes moderate salt-bridge interactions with Asp75 (58% occupancy) of NS3 and Asp83* (58% occupancy) of NS2B and a moderate hydrogen bond with Asp83* (65% occupancy) of NS2B. Its backbone also makes a strong hydrogen bond with Gly151 (99% occupancy) of NS3 (**Figure 2A**). We also noted that the P3 K of YKRK makes a moderate hydrogen bond with Phe84* (72% occupancy) of NS2B and is favorably placed in the S3 site to facilitate long-range electrostatic interactions with Asp75 (NS3), Asp79* (NS2B), and Asp83* (NS2B). Although P2 R in YRRR can make a

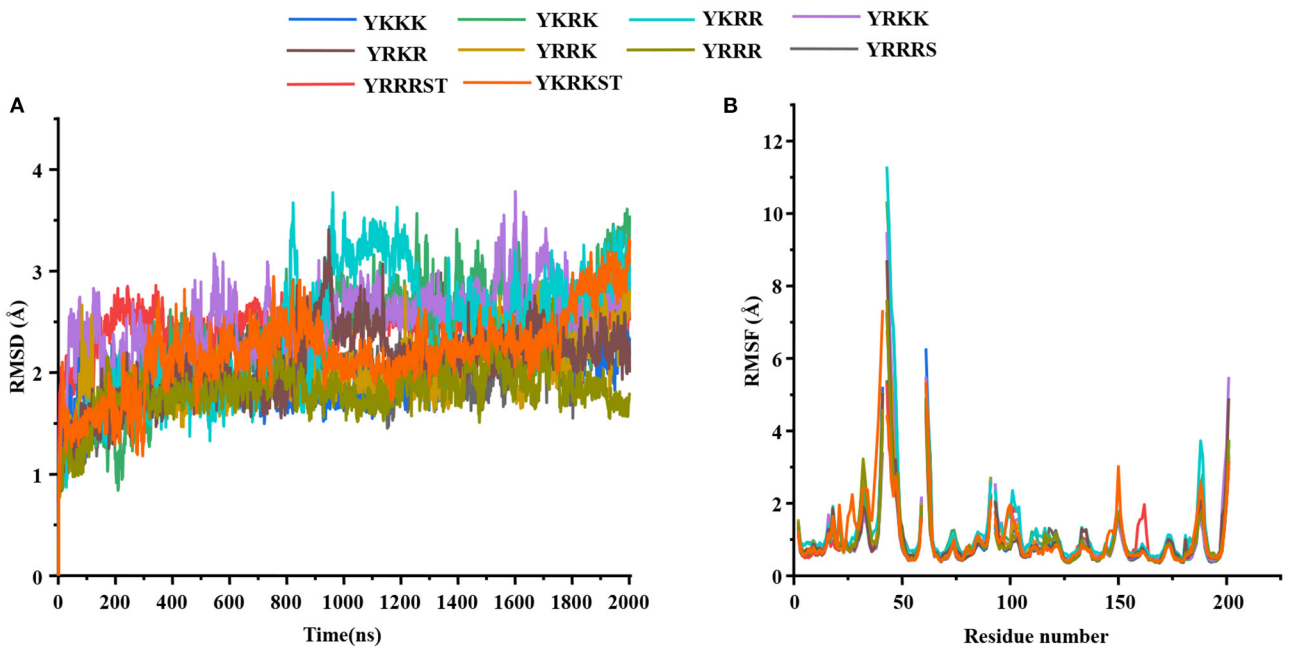


FIGURE 1 | (A) The root mean square deviations (RMSD) (Å) of the C_α atoms of the protease and **(B)** the root mean square fluctuations (RMSF) (Å) of different residues of the protease in different protein-peptide complexes.

TABLE 1 | The relative binding free energies (ΔG_{bind}) of different complexes.

Complex	ΔG_{bind} (kcal/mol)
YKKR-protease	−72.88 ± 8.15[a]
YKKK-protease	−59.16 ± 6.69
YKRK-protease	−80.52 ± 6.84
YKRR-protease	−62.47 ± 8.25
YRKK-protease	−72.87 ± 7.25
YRKR-protease	−56.33 ± 7.92
YRRK-protease	−74.02 ± 9.08
YRRR-protease	−77.94 ± 6.90
YRRRS-protease	−84.84 ± 8.46
YRRRST-protease	−107.49 ± 10.40
YKRKST-protease	−96.84 ± 9.21

[a]*Ref. (35).*

moderate salt-bridge interaction with Asp83* (65% occupancy) of NS2B, it failed to make an ionic interaction with Asp75 of NS3 (**Supplementary Figure S7A**). However, it can make moderate and weak hydrogen bonds with Asp75 (76% occupancy) of NS3 and Ser81* (37% occupancy) of NS2B, respectively (**Figure 2B**). Remarkably, the P3 R in YRRR failed to interact with the NS2B residues (**Figure 2B**; **Supplementary Figure S7A**). These results suggest that the binding of YRRR to the protease is weaker compared to YKRK.

Interestingly, if we compare the ΔG_{bind} of YKKR-protease and YRRR-protease complexes, it is clear that the latter complex is about 5 kcal/mol more stable than the former. This is in agreement with an earlier study (30) where a slightly higher IC_{50} value was obtained for a peptidomimetic inhibitor that contained RRR compared to KKR. It should be mentioned that as both YKKR and YRRR contain P1 R but differ at the P2 residue and the P3 residue does not make any direct ionic interaction with the protease, the higher stability of the latter is mainly due to the stronger binding of P2 R with the residues of S2 site. These results indicate that an ionic peptide with K at the P1 position and R at the P2 position may yield better potency. Also, the higher ΔG_{bind} of YKRK suggests that K would be preferred at the P3 position. In an earlier kinetic study for ZIKV (37), K was found to be preferred at the P3 position, while neutral residues, such as Trp, Tyr, Asp, and Pro at the same position showed no substrate activity. Similarly, in all stereotypes of the DENV protease, K/R at P1, R at P2, and K at P3 position were found to be highly favored (36). The yellow fever virus also prefers K/R, R, and K at P1, P2, and P3 positions, respectively (37). However, for the WNV protease, K/R at the P1, K at P2, and K at the P3 position were found to be preferred (37, 38).

It should be mentioned that the bindings of different pentapeptide substrates, such as (1) acyl-Norleucine-Lysine-Lysine-Arginine-7-amino-4-carbamoyl-methyl coumarin (Ac-nKKR-ACC), (2) acyl-D-Arginine-Lysine-Ornithine-Arginine-7-amino-4-carbamoyl-methyl coumarin (Ac-D-RKOR-ACC), (3) acyl-D-Lysine-Lysine-Ornithine-Arginine-7-amino-4-carbamoyl-methyl coumarin (Ac-D-KKOR-ACC), and (4) benzoyl-Norleucine-Lysine-Arginine-Arginine-aminomethyl coumarin (Bz-nKRR-AMC) to the ZIKV protease were shown to produce stable complexes of ΔG_{bind} lying between −20.42 ± 5.26 and −43.53 ± 0.97 kcal/mol (without entropy calculations)

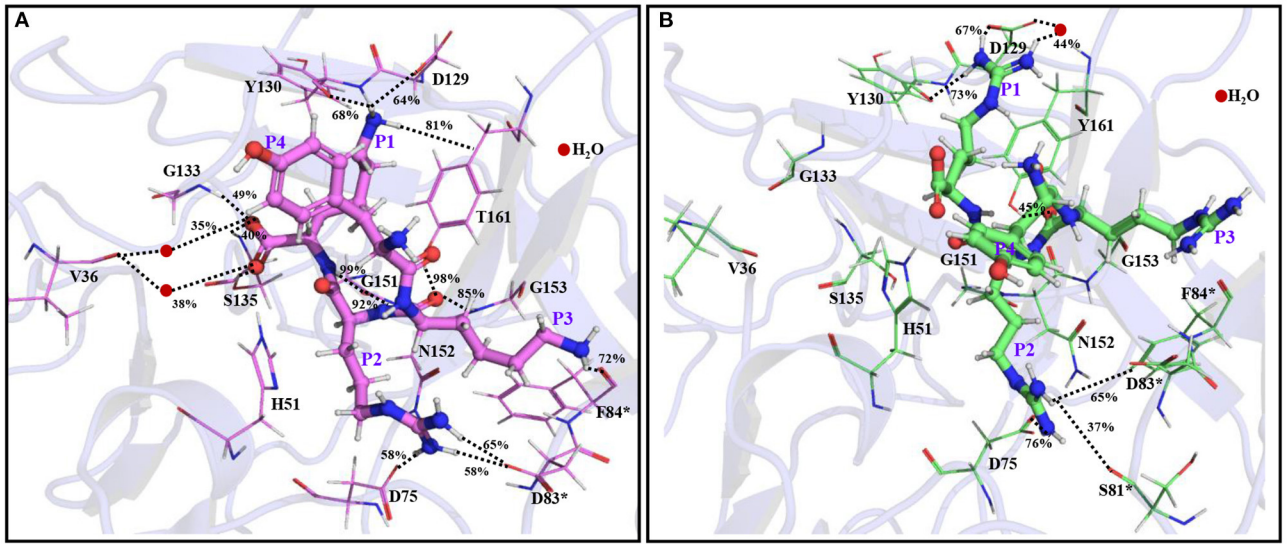


FIGURE 2 | The average simulated structures of **(A)** YKRK-protease and **(B)** YRRR-protease complexes. The residues of NS2B are marked by *. The hydrogen bonding and ionic interactions are shown by dotted lines. The percentage occupations of different interactions are also mentioned. The P1-P4 residues of the inhibitors are labeled.

(52). Among these substrates, the Ac-D-RKOR-ACC-protease complex was found to be the most stable one (52). This result led the authors to propose that R and O are preferred at P1 and P2 positions, respectively. However, as the study was not undertaken by considering P1 K and P2 K, the results obtained therein are not conclusive (52). Further, as P1 R and P2 O in the Ac-D-RKOR (substrate 2) were not making key ionic interactions with Asp129 and Asp83* of S1 and S2 sites, respectively, the Ac-D-RKOR-protease complex will be less stable than that of the YKRK-protease complex (52). Although the ΔG_{bind} values of the above substrate-protease complexes were obtained by using the MM/PBSA technique and the AMBER16 force field (52), the results cannot be directly correlated with the results obtained herein. However, the binding modes of the substrates/inhibitors and the relative difference in their ΔG_{bind} values support the higher stability of YKRK compared to Ac-D-RKOR.

The C-Terminal Extension of YRRR

As the C-terminal carboxyl group of some of the tetrapeptides including YKRK makes indirect hydrogen bonds with residues of prime sites, the C-terminal extension of these peptides may likely generate additional interactions with the residues of S1' and S2' sites. For these reasons, the two most stable peptides (YKRK and YRRR) were extended by adding one or two residues at their C-terminal. Initially, the YRRR was extended by adding S to its C-terminal (P1' position). Subsequently, the YRRRS peptide was further extended by adding T at the P2' position. These extended peptides can behave as the substrate competitive inhibitors and fully occupy the binding site of the protease.

It is found that the YRRRS-protease complex is about 7 kcal/mol more stable than that of the YRRR-protease complex and the YRRRST-protease complex is about 23 kcal/mol more stable than the YRRRS-protease complex (**Table 1**). This indicates that the C-terminal extension of YRRR by P1' S and P2' T will enhance its stability by 30 kcal/mol. Interestingly, the positively charged guanidine group of P1 R in YRRRS rotated 180 degrees from its initial conformation (YRRR) to make a strong ionic interaction with the Asp129 (85% occupancy) (**Figure 3B**, **Supplementary Figure S7**). Further, in this conformation, it makes several direct and indirect interactions of moderate stability with Asp129, Tyr130, and Tyr161 of NS3 (**Supplementary Figure S7B**). The P2 R makes weak interactions with Asp75 of NS3 and Asp83* of NS2B (<50% occupancy) and P3 R makes several interactions (<50% occupancy) with Asp79* and Phe84* of NS2B (**Figure 3A**, **Supplementary Figure S7B**). Notably, the ionic interaction between P3 R and Asp79* of NS2B is missing in the YRRR-protease complex (**Figure 2B**). Moreover, the N-terminal Y moved up toward the C-terminal S and its terminal NH_3^+ group makes a weak electrostatic interaction with Asp83* of NS2B (31% occupancy). Other than these, the backbones of P1' S, P2 R, P3 R, and P4 Y are found to make moderate hydrogen bonds with Gly133, Ser135, Gly153, Tyr161, and Asp83* (**Figure 3A**). Although only the backbone interaction of P1' S with Gly133 (80% occupancy) is obtained, its inclusion at the C-terminal helped P1 R to make stronger interactions with the residues of the S1 site, in particular with Asp129. These results suggest that YRRRS would make stronger interactions with the protease compared to YRRR.

Remarkably, further extension of YRRRS to YRRRST abolished interactions of P1 R with Asp129, Tyr130, and Y161 (**Figure 4A**; **Supplementary Figure S9A**) as the loop containing these residues moved away from the inhibitor. However, it made new contacts (hydrogen bonds) with Ala132 (61%

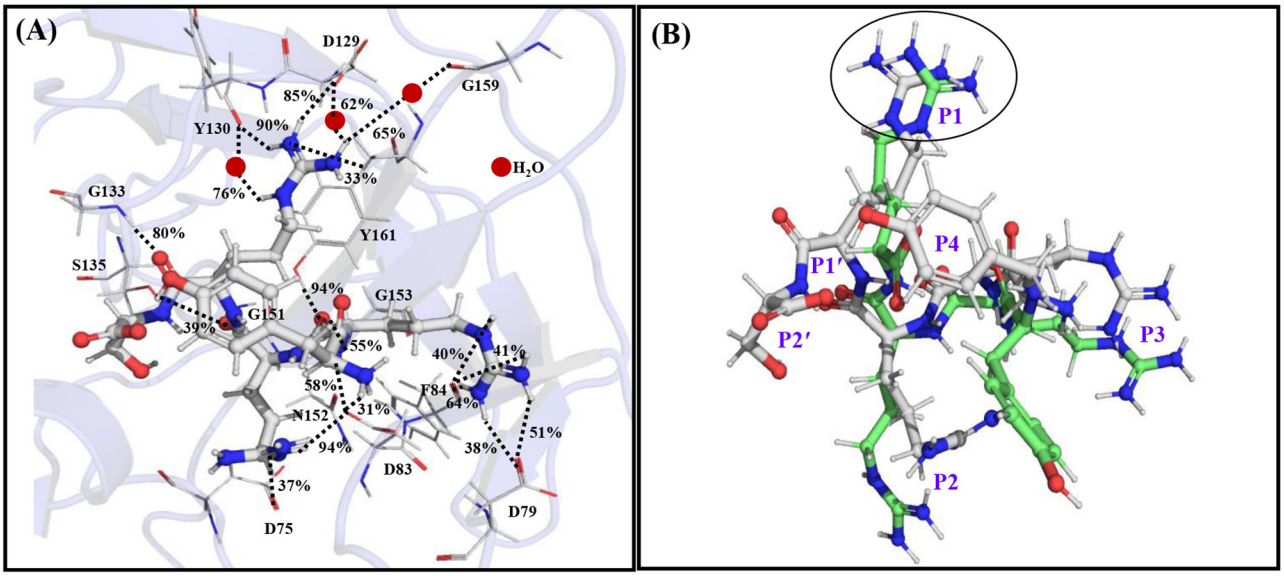


FIGURE 3 | (A) The average simulated structure of the YRRRS-protease complex. The different interactions between the peptide and protease are shown by dotted lines and the percentage occupations of these interactions evolved throughout the simulation are also mentioned. **(B)** The superimposition of the average simulated structure of the YRRRS-protease [**(C)** atoms are shown in white] onto the average simulated structure of the YRRR-protease [**(C)** atoms are shown in green] complex (protease structures are not shown). The P2'-P4 residues of the inhibitors are labeled in **(B)**.

occupancy) and Tyr134 (56% occupancy) and maintained the same orientation of the guanidine group as is obtained in the case of YRRRS (**Figure 3A**; **Supplementary Figure S7B**). In this structural arrangement, the side-chain hydroxyl group of P1' S makes a direct but weak hydrogen bond with His51 (48% occupancy) and a water-mediated moderate hydrogen bond with Asp75 (67% occupancy). Its carbonyl backbone also makes a water-mediated moderate hydrogen bond with Val36 (79% occupancy) (**Figure 4A**; **Supplementary Figure S9A**). The P2' T is found to make stable direct hydrogen bonds with Gly133 (83% occupancy) and Gly136 (88% occupancy) and a moderate hydrogen bond with Ser135 (64% occupancy). It also makes stable indirect water-mediated hydrogen bonds with Gln35 (80% occupancy). Interestingly, the interactions of P2 R with the residues of the S2 site improved significantly. It is now making a moderate ionic interaction with Asp75 of NS3 (74% occupancy) and strong hydrogen bonds with Asp83* of NS2B (89% occupancy) and Asp75 of NS3 (94% occupancy). It also makes moderate and weak hydrogen bonds with Ser81* (70% occupancy), and Lys73* (38% occupancy) of NS2B, respectively (**Figure 4A**; **Supplementary Figure S9A**). Other than these, the backbone atoms of P2 R and P3 R are making direct and indirect hydrogen bonds with Asn152, Asp75, and Gly153 of NS2B. However, the side-chain of P3 R and P4Y do not make any contact with the protease. Hence, the tight binding of the inhibitor with the prime site residues may have contributed to the higher stability of the YRRRST-protease complex.

The C-Terminal Extension of YKRK

As the binding of YRRRST hexapeptide to the protease is more stable than that of YRRRS, we added S and T to the C-terminal of YKRK to generate YKRKST-protease complex (**Figure 4B**; **Supplementary Figure S9B**). It is found that the YKRKST-protease complex is about 16 kcal/mol more stable than the YKRK-protease complex (**Table 1**). However, it is about 11 kcal/mol less stable than the YRRRST-protease complex (**Table 1**). This is a bit surprising and indicates that the preferences of K, R, and K at the P1, P2, and P3 positions, respectively get lost after the inclusion of P1' and P2' residues. Interestingly, a moderate hydrogen bond between the P1' S and Val36 (73% occupancy) is the only interaction found between the inhibitor and the prime site residues (**Figure 4B**; **Supplementary Figure S9B**). However, the P1 K was placed in the S1 site and is making a moderate electrostatic interaction with Asp129 (73% occupancy), a moderate hydrogen bond with Tyr130 (70% occupancy), and a strong π-cation interaction with Tyr161 (98% occupancy) (**Figure 4B**). Its carbonyl backbone is also making strong hydrogen bonds with Gly133 (98% occupancy) and Ser135 (83% occupancy) of NS3 and amide backbone is making a strong hydrogen bond with Gly151 (98% occupancy) of NS3. This clearly shows that the binding of P1 K to the S1 site remains to be strong despite the C-terminal extension. Although the side chain of P2 R makes an ionic and a hydrogen bond each with Asp75 (>50% occupancy) of NS3 it could not interact with Asp83* of NS2B. Nevertheless, its amide backbone makes a strong hydrogen bond with Gly151 (98% occupancy) of NS3 and its carbonyl backbone makes two hydrogen bonds one each with Gly153 (70% occupancy) and Tyr161 (99% occupancy) of NS3 (**Figure 4B**; **Supplementary Figure S9B**). Interestingly, the P3 K moved toward Asp83* of NS2B and is making a weak ionic interaction with it (46% occupancy). However, as F84* of NS2B rotated away from P3 K, it could not interact with F84*.

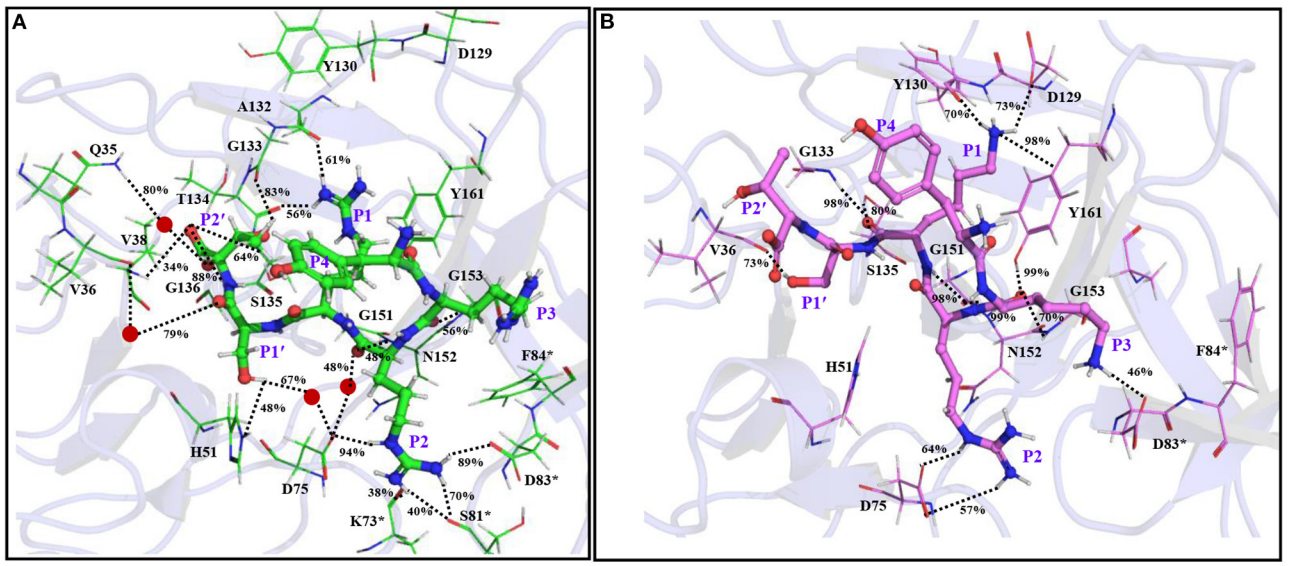


FIGURE 4 | The average simulated structures of the **(A)** YRRRST-protease (carbon atoms in green) and **(B)** YRKRST-protease (carbon atoms in violet) complexes. Different interactions are shown by dotted lines and their percentage occupations evolved during the simulations are also mentioned. Residues of NS2B are marked by *. The P1'–P4 residues of the inhibitors are labeled.

These interaction profiles indicate that although YKRKST makes stable interactions with the residues of non-prime sites (S1–S3), it failed to interact strongly with the prime site residues (S1' and S2' sites). This likely is the cause of its lower stability than that of YRRRST.

Effect of the Protein Dynamics

The superimposition of the average structures of YRRRST and YKRKST bound to NS2B–NS3 protease of ZIKV indicates that the loop regions of the protease containing residues from S1, S2, and S3 sites move significantly during the simulations. In the case of the YRRRST-protease complex, the movement of the S1 site away from the P1 R is significant compared to that of the YKRKST-protease complex (**Figures 4**, **5**; **Supplementary Figure S9**). Due to this reason, the P1 R of YRRRST is weakly bound to the S1 site (**Figure 5A**). Although the loop containing K73, D75, etc. of NS2B that partly constitutes the S2 site does not move much, the loop containing S81*, D83*, F84*, etc. of NS2B moves significantly in the YRRRST-protease complex. As was observed earlier (35) and is evident from **Figure 5A**, His51 is quite flexible and can adopt different conformations as per the ligand orientation (**Figure 5A**). This structural reorientation in the YRRRST-protease complex created a wide and deep S1 pocket (**Supplementary Figure S12**). Similarly, a wide pocket is also found at the prime site (**Supplementary Figure S12**).

Interestingly, except for the C-terminal primed residues (P1' and P2'), non-primed residues (P1–P3) of YRRRST and YKRKST adopted identical conformations. The conformation of P4 Y is slightly different in these two peptides as it is exposed to solvent and therefore, enjoys conformational flexibility. If we compare the conformation of YRRRST and YKRKST with the X-ray conformation of a peptidomimetic inhibitor-containing KKR (PDB ID 5ZMQ) (30), it is clear that P1, P2, and P3 residues of all of these inhibitors adopt an identical conformation (**Figures 5B,C**). However, the structural similarity of YKRKST with the peptidomimetic inhibitor is more rigorous (**Figure 5C**). These results indicate that the hexapeptides would act as potent inhibitors of protease activities. It also indicates that the preference of basic residues, such as K, R, and K at P1, P2, and P3 positions are not necessary if the peptide inhibitor is extended at the C-terminal. Nevertheless, the presence of basic residues (K/R) at the P1–P3 positions is required to get optimal inhibitory activities. Due to structural similarities between ZIKV, WNV, and DENV, it is expected that these inhibitors may act as pan-flavivirus inhibitors.

It should be mentioned that as the protease cleaves the substrate by cleaving the P1-P1' CN bond (**Supplementary Figure S11**) with the help of Ser135, His51, and Asp75 of NS3, the hexapeptides studied herein may be converted to tetrapeptides (30, 53). We found that in the case of YRRRST, the Ser135 and Gly133 oxyanion hole are moved away from the carbonyl O of the P1 residue and hence cannot make any hydrogen bond with it (**Supplementary Figure S11A**). This type of inactive conformation of oxyanion hole was reported earlier (54). However, the same oxyanion hole is active in the case of YKRKST as evident by two strong hydrogen bonds made by Gly133 and Ser135 with the carbonyl O of the P1 residue of the YKRKST (**Supplementary Figure S11B**). Interestingly, as the C (P1 residue)-O (Ser135) bond distance is almost identical (3.4 Å in YKRKST and 3.6 Å ion YRRRST) in these two complexes (**Supplementary Figure S11**), the

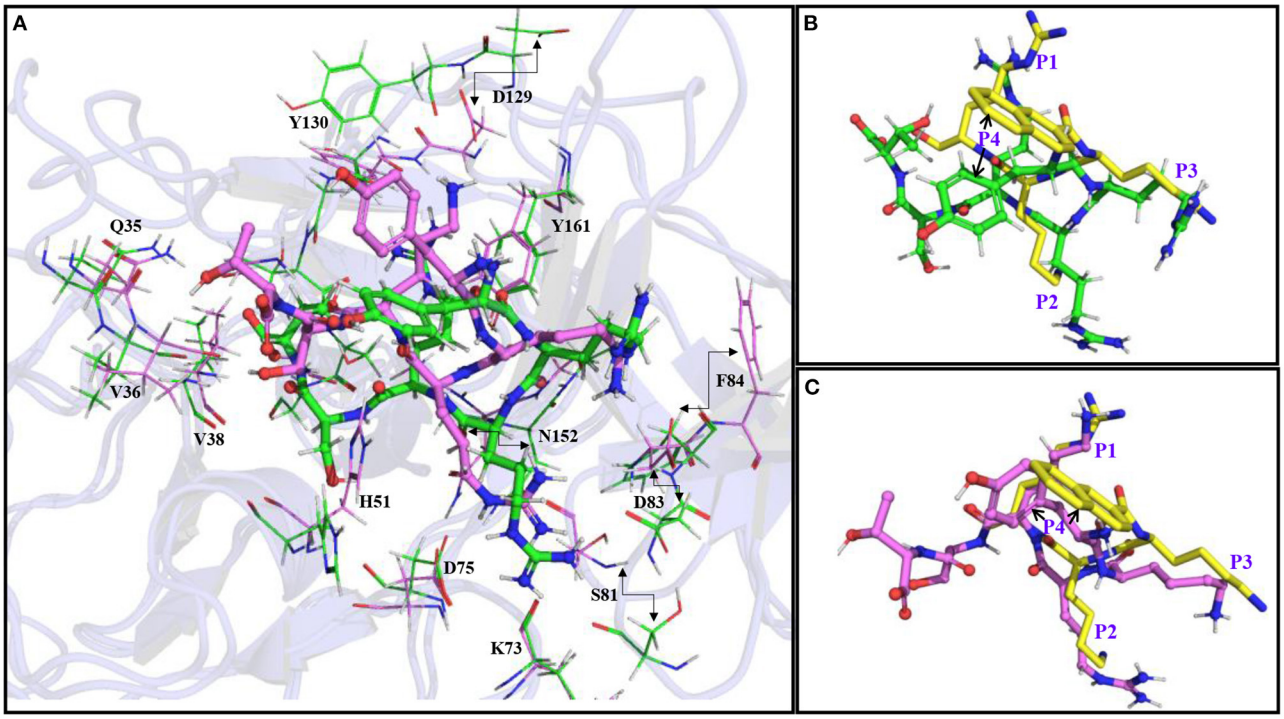


FIGURE 5 | (A) Superimposition of the average simulated structure of the YRRRST-protease complex (carbon atoms are shown in green) onto the average simulated structure of the YKRKST-protease complex (carbon atoms are shown in violet). **(B)** Superimpositions of the average simulated structures of the YRRRST-protease complex and **(C)** YKRKST-protease complex onto the peptidomimetic–protease complex containing KKR (PDB ID 5ZMQ) (carbon atoms in yellow). The movements of some residues during the simulations are indicated by arrows. The P1–P4 residues of the inhibitors are labeled in **(B,C)**.

Ser135 may get covalently bonded to the P1 carbonyl, thereby eventually leading to the scission of the P1-P1' peptide bond (53). Nevertheless, the results obtained herein highlight the importance of extended inhibitors in fully occupying the substrate-binding site of the protease. As residues beyond S2' site participate in the membrane-binding (36), further extension of the hexapeptide inhibitors may not yield encouraging results. Therefore, it is likely that hexapeptide inhibitors that contain an unnatural backbone may not be cleaved by the protease and hence would be extremely useful in inhibiting protease activities. Further, as in all of the inhibitors studied herein, the N-terminal P4 Y extends toward the C-terminal residue; cyclic inhibitors that connect C- and N-terminals (32, 33) may also be useful in inhibiting the protease activities. Similarly, the use of peptidomimetics or small molecular ligands that can produce equivalent interactions as generated by basic residues and are capable of crossing the hydrophobic lipid bilayers would be highly beneficial.

CONCLUSION

The present study revealed that the binding of a tetrapeptide inhibitor to the NS2B–NS3 protease of the ZIKV would create a highly stable complex ($\Delta G_{bind} = \sim -81$ kcal/mol) when the P1, P2, and P3 positions of the inhibitor contain K, R, and K, respectively. As the P4 position is exposed to solvent and acts as an anchoring group to stabilize the peptide in the substrate-binding site, Y at this position would serve the purpose. The presence of consecutive Rs at P1, P2, and P3 positions and Y at the P4 position would create the second most stable peptide-protease complex ($\Delta G_{bind} = \sim -78$ kcal/mol). However, the consideration of other combinations of K and R at these positions would also generate potent inhibitors as the ΔG_{bind} of such inhibitor-protease complexes lie between ∼-59 and −74 kcal/mol. The C-terminal extension of YRRR and YKRK by including P1' S and P2' T would generate even more stable complexes with $\Delta G_{bind} = \sim$-107 kcal/mol and ∼ −97 kcal/mol, respectively. These hexapeptides would behave as better substrate-competitive inhibitors than those of the tetrapeptide inhibitors. Thus, this study has indicated that there is enormous potential to develop potent inhibitors of the ZIKV protease by modifying the C-terminal prime residues with unnatural backbones or unnatural amino acids.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Materials**, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

NRJ conceived the research, analyzed the results, and wrote the paper. SP performed simulations. NRJ and SP approved the paper. Both authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

NRJ is thankful to the Council of Scientific and Industrial Research (CSIR, New Delhi) for financial assistance [Grant No-01(3061)21/EMR II]. SP is thankful to the department of pharmaceuticals and NIPER, Kolkata for providing a Senior Research Fellowship. NRJ and SP are thankful to the funding agency, IIITDM, and NIPER Kolkata for necessary computational facilities.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmed.2022.921060/full#supplementary-material

REFERENCES

1. Sharma V, Sharma M, Dhull D, Sharma Y, Kaushik S, Kaushik S. Zika virus: an emerging challenge to public health worldwide. *Can J Microbiol.* (2020) 66:87–98. doi: 10.1139/cjm-2019-0331
2. Agumadu VC, Ramphul K. Zika virus: a review of literature. *Cureus.* (2018) 10:e3025. doi: 10.7759/cureus.3025
3. Parra B, Lizarazo J, Jiménez-Arango JA, Zea-Vera AF, González-Manrique G, Vargas J, et al. Guillain–Barré syndrome associated with Zika virus infection in Colombia. *N Engl J Med.* (2016) 375:1513–23. doi: 10.1056/NEJMoa1605564
4. Mecharles S, Herrmann C, Poullain P, Tran TH, Deschamps N, Mathon G, et al. Acute myelitis due to Zika virus infection. *Lancet.* (2016) 387:1481. doi: 10.1016/S0140-6736(16)00644-9
5. Besnard M, Eyrolle-Guignot D, Guillemette-Artur P, Lastere S, Bost-Bezeaud F, Marcelis L, et al. Congenital cerebral malformations and dysfunction in fetuses and newborns following the 2013 to 2014 Zika virus epidemic in French Polynesia. *Euro Surveill.* (2016) 21. doi: 10.2807/1560-7917.ES.2016.21.13.30181
6. Antoniou E, Orovou E, Sarella A, Iliadou M, Rigas N, Palaska E, et al. Zika virus and the risk of developing microcephaly in infants: a systematic review. *Int J Environ Res Public Health*. (2020) 17:3806. doi: 10.3390/ijerph17113806
7. Chibueze EC, Tirado V, Lopes KD, Balogun OO, Takemoto Y, Swa T, et al. Zika virus infection in pregnancy: a systematic review of disease course and complications. *Reprod Health.* (2017) 14:28. doi: 10.1186/s12978-017-0285-6
8. Wen Z, Song H, Ming GL. How does Zika virus cause microcephaly? *Genes Dev.* (2017) 31:849–61. doi: 10.1101/gad.298216.117
9. Chen Q, Gouilly J, Ferrat YJ, Espino A, Glaziou Q, Cartron G, et al. Metabolic reprogramming by Zika virus provokes inflammation in human placenta. *Nat Commun.* (2020) 11:2967. doi: 10.1038/s41467-020-16754-z
10. Gulland A. Zika virus is a global public health emergency, declares WHO. *BMJ.* (2016) 352:i657. doi: 10.1136/bmj.i657
11. Bernatchez JA, Tran LT, Li J, Luan Y, Siqueira-Neto JL, Li R. Drugs for the treatment of Zika virus infection. *J Med Chem.* (2020) 63:470–89. doi: 10.1021/acs.jmedchem.9b00775
12. Morabito KM, Graham BS. Zika virus vaccine development. *J Infect Dis.* (2017) 216:S957–63. doi: 10.1093/infdis/jix464
13. Kumar A, Liang B, Aarthy M, Singh SK, Garg N, Mysorekar IU, et al. Hydroxychloroquine inhibits Zika virus NS2B-NS3 protease. *ACS Omega.* (2018) 3:18132–41. doi: 10.1021/acsomega.8b01002
14. Sirohi D, Kuhn RJ. Zika virus structure, maturation, and receptors. *J Infect Dis.* (2017) 216:S935–44. doi: 10.1093/infdis/jix515
15. Wahaab A, Mustafa BE, Hameed M, Stevenson NJ, Anwar MN, Liu K, et al. Potential role of flavivirus NS2B-NS3 proteases in viral pathogenesis and anti-flavivirus drug discovery employing animal cells and models: a review. *Viruses.* (2022) 14:44. doi: 10.3390/v14010044
16. Hilgenfeld R, Lei J, Zhang L. The structure of the Zika virus protease, NS2B/NS3(pro). *Adv Exp Med Biol.* (2018) 1062:131–45. doi: 10.1007/978-981-10-8727-1_10
17. Phoo WW, Li Y, Zhang Z, Lee MY, Loh YR, Tan YB, et al. Structure of the NS2B-NS3 protease from Zika virus after self-cleavage. *Nat Commun.* (2016) 7:13410. doi: 10.1038/ncomms13410
18. Hill ME, Kumar A, Wells JA, Hobman TC, Julien O, Hardy JA. The unique cofactor region of Zika Virus NS2B-NS3 Protease facilitates cleavage of key host proteins. *ACS Chem Biol.* (2018) 9:2398–405. doi: 10.1021/acschembio.8b00508
19. Voss S, Nitsche C. Inhibitors of the Zika virus protease NS2B-NS3. *Bioorg Med Chem Lett.* (2020) 30:126965. doi: 10.1016/j.bmcl.2020.126965
20. Pathak N, Kuo YP, Chang TY, Huang CT, Hung HC, Hsu JT, et al. Zika Virus NS3 protease pharmacophore anchor model and drug discovery. *Sci Rep.* (2020) 10:8929. doi: 10.1038/s41598-020-65489-w
21. Kang C, Keller TH, Luo D. Zika virus protease: an antiviral drug target. *Trends Microbiol.* (2017) 25:797–808. doi: 10.1016/j.tim.2017.07.001
22. Brecher M, Zhang J, Li H. The flavivirus protease as a target for drug discovery. *Virol Sin.* (2013) 28:326–36. doi: 10.1007/s12250-013-3390-x
23. Li Y, Zhang Z, Phoo WW, Loh YR, Li R, Yang HY, et al. Structural insights into the inhibition of Zika virus NS2B-NS3 protease by a small-molecule inhibitor. *Structure.* (2018) 26:555–64 e3 doi: 10.1016/j.str.2018.02.005
24. Poulsen A, Kang C, Keller TH. Drug design for flavivirus proteases: what are we missing? *Curr Pharm Des.* (2014) 20:3422–7. doi: 10.2174/13816128113199990633
25. Li Q, Kang C. Structure and dynamics of Zika virus protease and its insights into inhibitor design. *Biomedicines.* (2021) 9:1044. doi: 10.3390/biomedicines9081044
26. Roy A, Lim L, Srivastava S, Lu Y, Song J. Solution conformations of Zika NS2B-NS3pro and its inhibition by natural products from edible plants. *PLoS ONE.* (2017) 12:e0180632. doi: 10.1371/journal.pone.0180632
27. Shiryaev SA, Farhy C, Pinto A, Huang CT, Simonetti N, Elong Ngono A, et al. Characterization of the Zika virus two-component NS2B-NS3 protease and structure-assisted identification of allosteric small-molecule antagonists. *Antiviral Res.* (2017) 143:218–29. doi: 10.1016/j.antiviral.2017.04.015
28. Lei J, Hansen G, Nitsche C, Klein CD, Zhang L, Hilgenfeld R. Crystal structure of Zika virus NS2B-NS3 protease in complex with a boronate inhibitor. *Science.* (2016) 353:503–5. doi: 10.1126/science.aag2419
29. Li Y, Zhang Z, Phoo WW, Loh YR, Wang W, Liu S, et al. Structural dynamics of Zika virus NS2B-NS3 protease binding to dipeptide inhibitors. *Structure.* (2017) 25:1242–50 e3. doi: 10.1016/j.str.2017.06.006
30. Phoo WW, Zhang Z, Wirawan M, Chew EJC, Chew ABL, Kouretova J, et al. Structures of Zika virus NS2B-NS3 protease in complex with peptidomimetic inhibitors. *Antiviral Res.* (2018) 160:17–24. doi: 10.1016/j.antiviral.2018.10.006
31. Campos DMO, Katyanna S, Bezerra Esmaile SC, Fulco UL, Albuquerque EL, Oliveira JI. Intermolecular interactions of cn-716 and acyl-KR-aldehyde dipeptide inhibitors against Zika virus. *Phys Chem Chem Phys.* (2020) 22:15683–95. doi: 10.1039/D0CP02254C
32. Nitsche C, Passioura T, Varava P, Mahawaththa MC, Leuthold MM, Klein CD, et al. De novo discovery of nonstandard macrocyclic peptides as noncompetitive inhibitors of the Zika virus NS2B-NS3 protease. *ACS Med Chem Lett.* (2019) 10:168–74. doi: 10.1021/acsmedchemlett.8b00535
33. Braun NJ, Quek JP, Huber S, Kouretova J, Rogge D, Lang-Henkel H, et al. Structure-based macrocyclization of substrate analogue NS2B-NS3 protease inhibitors of Zika, West Nile and dengue viruses. *Chem Med Chem.* (2020) 15:1439–52. doi: 10.1002/cmdc.202000237

34. Li Y, Loh YR, Hung AW, Kang CB. Characterization of molecular interactions between Zika virus protease and peptides derived from the C-terminus of NS2B. *Biochem Biophys Res Commun.* (2018) 503:691–6. doi: 10.1016/j.bbrc.2018.06.062
35. Pant S, Bhattacharya G, Jena NR. Structures and dynamics of peptide and peptidomimetic inhibitors bound to the NS2B-NS3 protease of the ZIKA virus. *J Biomol Struct Dyn.* (2022) 3:1–13. doi: 10.1080/07391102.2022.2045223
36. Li J, Lim SP, Beer D, Patel V, Wen D, Tumanut C, et al. Functional profiling of recombinant NS3 protease from all four sterotypes of Dengu virus using tetrapeptide and octapeptide substrate libraries. *J Biol Chem*. 280:28766–74. doi: 10.1074/jbc.M500588200
37. Gruba N, Martinez JIR, Grzywa R, Wysocka M, Skorenski M, Burmistrz M. et al. Substrate profiling of Zika virus NS2B-NS3 protease. *FEBS Letters.* (2016) 590:3459–68. doi: 10.1002/1873-3468.12443
38. Ang MJ Li Z, Lim HA, Ng FM, Then SW, Wee JL. et al. A P2 and P3 substrate specificity comparison between the Murray Valley encephalitis and West Nile Virus NS2B/NS3 protease using C-terminal agmatine dipeptides. *Peptides.* (2014) 52:49–52. doi: 10.1016/j.peptides.2013.12.002
39. Nitsche C. Proteases from dengue, West Nile and Zika viruses as drug targets. *Biophys. Rev.* (2019) 11:157–65. doi: 10.1007/s12551-019-00508-3
40. Zhang Z, Li Y, Loh YR, Phoo WW, Hung AW. Kang C,et al. Crystal structure of the unlinked NS2B-NS3 protease from Zika virus. *Science.* (2016) 354:1597–600. doi: 10.1126/science.aai9309
41. Schrödinger, LLC. *The PyMOL Molecular Graphics System, Version 1.2r3pre*.
42. Jena NR, Pant S, Srivastava HK. Artificially expanded genetic information systems (AEGISs) as potent inhibitors of the RNA-dependent RNA polymerase of SARS-CoV-2. *J Biomol Struct Dyn*. (2021) 1–17. doi: 10.1080/07391102.2021.1883112
43. Pant S. Jena NR. Inhibition of the RNA-dependent RNA polymerase of the SARS-CoV-2 by short peptide inhibitors, European. *J Pharm Sci.* (2021) 167:106012. doi: 10.1016/j.ejps.2021.106012
44. ACM. *Proceedings of the 2006 ACM/IEEE Conference on Supercomputing. Tampa, Florida*. (2006). Association for Computing Machinery: New York, NY.
45. Schrödinger Release 2021-4. *Maestro-Desmond Interoperability Tools.* Desmond Molecular Dynamics System, D.E. New York, NY: Shaw Research. Schrödinger, New York, NY (2021).
46. Lu C, Wu C, Ghoreishi D, Chen W, Wang L, Damm W, et al. OPLS4: improving force field accuracy on challenging regimes of chemical space. *J Chem Theory Comput.* (2021) 17:4291–300. doi: 10.1021/acs.jctc.1c00302
47. Zielkiewicz J. Structural properties of water: comparison of the SPC, SPCE, TIP4P, and TIP5P models of water. *J Chem Phys.* (2005) 123:104501. doi: 10.1063/1.2018637
48. Saputro DRS, Widyaningsih P. Limited memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) method for the parameter estimation on geographically weighted ordinal logistic regression model (GWOLR). *AIP Conf Proc.* (2017) 1868:040009. doi: 10.1063/1.4995124
49. Posch HA, Hoover WG, Vesely FJ. Canonical dynamics of the nose oscillator: stability, order, and chaos. *Phys Rev A Gen phys.* (1986) 33:4253–65. doi: 10.1103/PhysRevA.33.4253
50. Martyna GJ, Tobias DJ, Klein ML. Constant pressure molecular dynamics algorithms. *J Chem Phys.* (1994) 101:4177–89. doi: 10.1063/1.467468
51. Petersen HG. Accuracy and efficiency of the particle mesh Ewald method. *J Chem Phys.* (1995) 103:3668–79. doi: 10.1063/1.470043
52. Nutho B, Rungrotmongkol T. Binding recognition of substrates in NS2B/NS3 serine protease of zika virus revealed by molecular dynamics simulations. *J Mol Graph Model.* (2019) 92:227–35. doi: 10.1016/j.jmgm.2019.08.001
53. F da Silva-Júnior, E, Schirmeister T, de Araújo-Júnior JX. *Recent Advances in Inhibitors of Flavivirus NS2B-NS3 Protease From Dengue, Zika, and West Nile Viruses, Book: Vector-Borne Diseases and Treatment*. Cambridge, United Kingdom: Open Access eBooks (2018). p. 1–25.
54. Ren J, Lee H, Kotak A, Johnson ME. MD simulations reveal alternate conformations of the oxyanion hole in the Zika virus NS2B/NS3 protease. *Proteins*. (2019) 2019:1–10. doi: 10.1002/prot.25809

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.