In silico and in vitro study of peptidomimetic protease inhibitors against Zika and Tick-Borne Encephalitis Virus

Lucas Brock

Master Degree Project in Infection Biology, 30 credits. Spring 2019
Supervisors: Johan Lennerstrand & Dario Akaberi
Clinical Microbiology, Department of Medical Sciences, Uppsala University
Abstract
Flaviviruses are single stranded positive sense RNA viruses spread via arthropod vectors. Flaviviruses include a number of important human pathogens such as Dengue, West Nile, Zika and Tick-Borne Encephalitis virus. Currently there are no antiviral treatments available against any of these viruses, even though a large part of the global population is at risk of infection. A possible target for drug development is the highly conserved viral protease NS2B/NS3 which has already been successfully used to develop approved drugs against the related Hepatitis C virus protease. Recent advances in the production of the Zika NS2B/NS3 protease and the discovery of pan-Dengue and West Nile Virus peptidomimetic protease inhibitors has provided the basis for the further development of potential inhibitors. Using in silico molecular dynamics techniques, additional compounds were designed and then tested in an in vitro enzymatic assay against the Zika NS2B/NS3 protease to better understand the mechanism of action of these inhibitors. It was found that a previously reported tripeptide protease inhibitor against Dengue and West Nile Virus also had a sub micro-molar IC\textsubscript{50} against Zika. Additionally, a novel dipeptide protease inhibitor was found to have a sub micro-molar IC\textsubscript{50} against Zika. This paves the way for the further development of a more potent drug-like compound, and eventually a possible pan-Flavivirus protease inhibitor.

Killing two birds with one stone: in search of a pan-Flavivirus treatment (Popular scientific summary)
The genus of Flavivirus is large and contains several human pathogens, such as Dengue and Zika. These viruses are spread by ticks or mosquitoes and are found from the Baltic coast to the jungles of Brazil. Despite a significant portion of the global population is at risk of infections that have the potential to be crippling or life threatening, there are no antiviral treatments available against any member of this genus. Fortunately, these viruses all share a relatively well conserved protease, named NS2B/NS3, which is essential to the life cycle of the virus. The image below shows the protease of Dengue (in red) and West Nile Virus (in yellow) compared to the protease of Zika (in green), demonstrating the similar structures of these different proteases.

The similarity of the proteases is also shown by the recent discovery of pan-Dengue and West Nile Virus inhibitors, based on the sequence of amino acids that the protease recognizes in the virus. The picture below shows the general structure of these inhibitors, with the amino acids circled in red.
Using variants of these inhibitors, we aimed to find out how they work by simulating their interactions with Zika and TBE protease. First, we fit the inhibitor with a different R group into the active site of the protease in a process called docking. The inhibitor can be docked in several different conformations, so the correct ‘pose’ was chosen based on previous knowledge. Once the correct pose was chosen, we used a super computer to simulate the inhibitor interacting with the protease for 50 nanoseconds in conditions that aim to mimic real life. The variants of the compounds tested this way were also produced in order to see how well they could inhibit Zika protease in real life. To this end we tested seven different compounds. Two compounds had been tested on Dengue and West Nile Virus and we found that they could inhibit the protease of Zika, although not as well. This was predicted by our simulations, which showed a more unstable interaction in Zika compared to other Flaviviruses. Another version we produced changed the R group to a histidine. This was done in order to figure out the correct binding site for the R group. As shown in the picture below, a previously published docking pose shows that the R group (circled in red) is at the S1 site but we found that this is unstable and that the R group actually interacts with the S1’ site (indicated by the green arrow).

If the R group really did bind to the S1 site, the version with a histidine should show an inhibitory concentration similar to the molecule from which it was derived, as it should be able to interact with the S1 site. This was not the case. Finally we also tested a new type of inhibitor, which was a dipeptide instead of a tripeptide. This type of molecule is more ‘drug-like’ and it showed a good ability to inhibit the protease. This means that we are on the right track in developing a potential inhibitor against a group of viruses that threatens large populations of the globe, and for which there is no treatment yet.
Introduction

**Flaviviruses** comprise a genus within the *Flaviviridae* family. Within this genus, more than 50% of the species are associated with human disease\(^1\), including important pathogens such as Yellow Fever virus (YFV), Dengue Virus (DEV), Zika virus and tick-borne encephalitis virus (TBEV). These *Flaviviruses* share several characteristics. They are enveloped single stranded positive sense RNA viruses, with the genome that is directly translated and encodes a single polyprotein. This polyprotein is processed by host proteases and the viral NS2B/NS3 protease to form three structural and seven non-structural proteins. *Flaviviruses* are distributed over the globe and can be divided into three distinct groups depending on their vector. Group 1 is spread by mosquitoes of the *Aedes* and *Culex* species, with *Flaviviruses* associated with *Aedes* mosquitoes causing haemorrhagic diseases, and those linked to *Culex* species generally causing encephalitis.\(^2\) Group 2 is spread by ticks and group 3 is spread by unknown vectors. Although they have a global distribution, many species are endemic to certain areas and infection by *Flaviviruses* can have a variety of symptoms depending on the species. Most infections are relatively mild, as is the case with most Zika infections in adults, but some can be very severe or cause severe complications, as is the case with TBE or severe Dengue fever.\(^3\)

There are currently no treatment options available against *Flaviviruses* but there are several effective vaccines available for Yellow Fever, Japanese encephalitis and TBE.\(^3\) Although effective, there are some limitations to these vaccines. For example the TBE vaccine requires several doses over a minimum period of 6 months\(^4\) and others have limited availability in non-endemic areas.\(^5\) For other *Flaviviruses*, such as Dengue and Zika, there are vaccines in development but none available as of now. It has been estimated that Dengue alone poses a risk to 3.97 billion people\(^6\) globally and therefore without any treatments available to supplement vaccines, a huge portion of people are at risk of Flavivirus infection, e.g. those living in endemic areas and travellers and workers wishing to go to there. A promising avenue of investigation is in the use of NS2B/NS3 inhibitors. As mentioned, it is responsible for the processing of the viral polyprotein through its chymotrypsin-like serine protease domain, as well as containing a C-terminal NTPase-Helicase that is involved in viral replication\(^7\), making it essential in the viral life cycle and therefore an attractive target.

Similar strategies have already shown to be effective in the treatment of Hepatitis C virus (HCV), belonging to the *Hepacivirus* genus of the *Flaviviridae* family\(^8\), with the first drugs approved against its NS3/NS4A protease in 2011. Furthermore, the NS2B/NS3 protease complex is highly conserved amongst different Flavivirus species and serotypes\(^9\). This degree of conservation could provide the basis of a pan-Flavivirus protease inhibitor that could be used in the treatment of a variety of species.

There are unfortunately some difficulties in developing inhibitors to these proteases. The Flavivirus proteases tend to recognise positively charged residues in the binding pocket at the P1 and P2 positions which reflects the native amino acid sequence it cleaves in the polyprotein. This can cause issues with bioavailability and cell membrane penetration for polar drug candidates. They also have a relatively shallow binding pocket,\(^10,11\) which might require the use of a large inhibitor. A large inhibitor might in turn have issues with solubility, as well as other pharmacological traits such as bioavailability and tissue distribution. However, knowledge about the target can greatly aid the drug discovery process, especially if the 3D structure of the target is known. The gold standard for this is crystal structures but until recently most crystal structures elucidated for Flavivirus proteases have been of a covalently linked NS2B/NS3, where NS2B is attached to NS3 via a glycine-serine-glycine
(G$_4$SG$_4$) linker. This does not reflect the native state of the protein, where NS2B forms a complex with NS3 to form the fully functional enzyme. While this form of the protease is active and soluble, it has been shown that, at least in the case of Zika and Dengue, the unlinked protease shows a higher activity than the linked version, which may be due to steric hindrance introduced by the linker region.\textsuperscript{12-14} Fortunately, a crystal structures has become available for Zika protease in its unlinked native state, which could reduce the detection of false positives and thus improve the screening and designing of protease inhibitors.\textsuperscript{14} Several compounds with good levels of inhibition against Dengue and West Nile Virus proteases that acted as substrate mimics have also recently been developed by a group in Germany.\textsuperscript{11} These compounds were based of the preferred cleavage site sequences of these proteases together with modified C- and N-terminal groups, indicating that potent dual inhibitors can be developed against different Flavivirus proteases, paving the way for further development of anti-Flaviviral compounds.\textsuperscript{10}

While \textit{in vitro} high throughput screening of compounds is the most common way of identifying new drug candidates, it is costly and unpractical for smaller companies and research teams. Through the use of crystal structures, compounds can also be investigated using \textit{in silico} methods such as docking and molecular dynamics simulations. Docking a compound of interest to a protein can give information on the predominant binding pose of that compound, which can help in ranking compounds and in hit optimization. However, this technique only gives a static pose which gives no information on how the compound interacts with the protein over a period of time. Molecular dynamics simulations aim to address this by simulating the interactions between compound and protein. These simulations require significant computing power but can provide much more information than docking techniques.

Through the use of \textit{in silico} molecular dynamic simulation methods, our research group evaluated analogues of published compounds to explore the binding mechanisms and potentially allowing for the identification of promising new hits without the need for synthesising and testing a large amount of compounds \textit{in vitro}. Compounds were tested \textit{in silico} on WNV, DEV and Zika in addition to a TBE NS2B/NS3 protease that had been constructed previously by this group using homology modelling. This technique is used when there is no crystal structure available for a protein. It aims to create a 3D atomic model of a protein through the use of the proteins’ amino acid sequence and the crystal structure of a homologous protein, in this case Zika, Dengue and WNV proteases. This is possible due to the fact that protein structures are well conserved, even if the amino acid sequences are not identical. In the case of TBE, its NS3 domain has more than a 40% identical sequence to other Flaviviruses NS3 domains. This is well above the 30% required to construct an accurate model (personal communications, Dario Akaberi). Several compounds examined in the \textit{in silico} screen were also synthesised. In this study, \textit{in vitro} inhibitory enzymatic assays were set up and several of the \textit{in silico} compounds that were examined and synthesised were tested against TBE and Zika proteases to determine the half maximum inhibitory concentration (IC$_{50}$).
Methodology

**In silico docking and molecular dynamics simulations**

Docking and molecular dynamics (MD) simulations were carried out using a method previously published by this group\(^{15}\). Briefly, 3D structures for ligands were generated using MolConverter v16.7.4.0 (by ChemAxon) and NS2B/NS3 crystal structures for Zika (ID: 5GPI) and Dengue (ID: 3U1I) were acquired from the Protein Data Bank. The structure for Tick-Borne Encephalitis NS2B/NS3 was generated using homology modelling (unpublished results). Ligands were docked to the protein using AutoDock Vina v1.1.2. The best binding poses were then selected manually for use in further MD simulations. GROMACS version 5.1.1 was used to carry out MD simulations and computational resources were provided by Uppsala Multidisciplinary Centre for Advanced Computational Science (UPPMAX) via the The Swedish National Infrastructure for Computing (SNIC). Adjusted from the previous method, AMBER99SB-ILDN force field parameters were assigned to the target and parameters from the GAFF force field were assigned to the ligand. The system was solvated and ions were added to reach a physiological concentration of 0.15 M. Energy minimization was carried out and this was followed by first a temperature and then a pressure equilibration step. MD simulations were then carried out for 50 ns in triplicate.

**Enzyme production and purification**

Enzyme production and purification was kindly carried out and provided by the Protein Science Facility at the Karolinska Institute, Stockholm. The method for both Zika and TBE was based off a previously published method for the production of unlinked Zika and TBE protease.\(^{12}\) Briefly, *E. coli* strain BL21(DE3)-T1R pRARE2 was transformed with a plasmid carrying the Zika or TBE NS2B and NS3 gene and cultured. Proteins were purified and analysed using SDS-PAGE. Zika enzyme was delivered in 0.5 or 1 mg lyophilized aliquots. Purification of TBE resulted in three distinct fractions of enzyme, which were delivered in 0.5 or 1 mg lyophilized aliquots.

**Activity assay**

Purification of enzyme resulted in one distinct Zika fraction and three distinct TBE fractions. The purified Zika NS2B/NS3 protease and the three fractions of purified TBE NS2B/NS3 protease were diluted in inhibition assay buffer (20 mM Tris, pH 8.5, 10% glycerol, 0.01% Triton X-100). Activity was measured using a benzol-Nle-Lys-Arg-Arg-aminomethylcoumarin substrate (Bachem, Switzerland) (AMC-substrate) dissolved in inhibition assay buffer to a stock concentration of 1 mM. The protease was diluted to final concentrations of 100, 10 and 1 µM and added to a Corning 96 Well plate with 20 µM of the AMC-substrate for a final volume of 100 µL. Readings were carried out in triplicate using a Tecan plate reader with an excitation wavelength of 380 nm and an emission wavelength of 460 nm. Measurements were made every 30 seconds for 15 minutes.

**Enzyme kinetics assay**

Zika NS2B/NS3 protease activity assay was carried out using the AMC-substrate. The AMC-substrate was diluted to a starting concentration of 150 µM and serially diluted 1:2 to a final concentration of 4.69 µM in inhibition assay buffer (20 mM Tris, pH 8.5, 10% glycerol, 0.01% Triton X-100). The substrate was added to 3 nM of the protease diluted in the same buffer in a Corning 96 Well plate to a final volume of 100 µL. Readings were carried out in triplicate using a Tecan plate reader (Tecan) with an excitation wavelength of 380 nm and an
emission wavelength of 460 nm. Measurements were taken at minute intervals for 30 minutes at 37°C. In order to convert RFU to moles of AMC, a standard curve of free AMC using a two-fold dilution ranging from 400 to 0.2 μM was made. An assay of AMC in concentrations of 5, 0.5 and 0 μM with five-fold dilutions of compound 104 ranging from 100 to 0.8 μM was made to determine the effect of the fluorescent compounds on the signal produced. Initial rates of reactions were analysed using GraphPad.

**Fluorometric inhibition assay**

Inhibition assays of Zika NS2B/NS3 protease were carried out using the same protocol for different inhibitors, except compound 21C. Inhibitors were produced and were kindly provided by the Department of Medicinal Chemistry, Uppsala University. Lyophilized aliquots were dissolved in distilled water to a stock concentration of 1 mM. A starting concentration of 100 μM was used and serially diluted 1:5 in inhibition assay buffer (20 mM Tris–HCl, pH 8.5, 10% glycerol, 0.01% Triton X-100) to a final concentration of 6.4 nM. 3 nM of protease was incubated with the inhibitor and the reaction was started by the addition of 15 μM AMC-substrate. Total volume was 100 μL. Compound 21C was dissolved in DMSO and final reaction concentrations contained 5% DMSO. Readings were carried out in triplicate using a Tecan plate reader with an excitation wavelength of 380 nm and an emission wavelength of 460 nm. Measurements were taken at minute intervals for 30 minutes at 37°C. Initial rates of reactions were analysed using GraphPad.
**Results**

In order to further develop antivirals against *Flavivirus* proteases, an *in silico* study was set up to investigate potential inhibitors against TBE, Zika, Dengue and West Nile virus based on previously published compounds. Several compounds investigated in the *in silico* study were synthesised and then tested. First, Zika and TBE enzymes were produced and characterized. An *in vitro* inhibitory assay was then set up and optimized. This was followed by testing of several tripeptide protease inhibitors, as well as a novel dipeptide compound.

**Enzyme kinetics**

The unlinked Zika and TBE NS2B/NS3 proteases constructs were tested for activity using Bz-Nle-Lys-Arg-Arg-aminomethylcoumarin as fluorogenic substrate (AMC-substrate). This molecule is based on the preferential cleavage sequence of Flavivirus proteases and upon cleavage releases free AMC. Free AMC can then be measured using a fluorescent assay, giving an indication of enzyme activity. Unlinked Zika NS2B/NS3 protease, which had been produced following a previously published method\(^\text{12}\), showed activity but the TBE NS2B/NS3 protease produced using this method did not (data not shown). TBE was therefore not used for the remaining experiments.

After testing for initial activity, the enzyme kinetics of the unlinked Zika NS2B/NS3 protease was investigated to see if they correspond to previously published data using a increasing amount of substrate concentrations. As shown in Figure 1, the Km of this protease was found to be 8.29 ± 3.57 µM, which agrees with the Km of 6.33 ± 2.41 µM of a previously published unlinked Zika NS2B/NS3 protease employing the same substrate\(^\text{14}\), indicating our enzyme was working correctly.

![Figure 1. Enzyme activity of Zika NS2B/NS3 protease.](image)

**Figure 1.** Enzyme activity of Zika NS2B/NS3 protease. Activity of the protease was measured using an AMC-substrate at concentrations from 0 to 150 µM. The dotted line represents K. The assay was carried out in duplicate and measurements were taken in triplicate. Standard deviations are represented by error bars and the SE of the Km is shown.
Molecular dynamics and inhibition assays
Once the enzyme kinetics of the Zika virus protease was determined, several different compounds were tested for their ability to inhibit *in vitro* the ability of the protease to cleave the AMC-substrate. These compounds were based on a tripeptide compound scaffold previously shown to inhibit Dengue, West Nile and Zika NS2B/NS3.11,13 These tripeptide compounds are based on the consensus sequence recognized by NS2B/NS3 with the addition of C- and N-terminal groups. Compound 86, shown in Figure 2, had been tested against Zika before and our assays found it to have the same IC$_{50}$ of 0.958 μM as was previously reported.13 Figure 2 also shows the results of the 50 ns molecular dynamics simulations carried out using compound 86. The simulations show that, over this 50 ns period, the arginine residue interacts with the S2 pocket and that the lysine residue interacts with the S1 pocket. The N-terminal thiophene group interacts with the S3 pocket while the C-terminal group interacts with the S1’ pocket.

![Figure 2](image_url)

**Figure 2.** Molecular dynamics simulation and *in vitro* enzymatic inhibition graph of compound 86 against Zika protease. A) 50 ns molecular dynamics simulation of compound 86 with Zika NS2B/NS3 protease shows arginine interacting with the S2 pocket, lysine interacting with the S1 pocket and the N-terminal and C-terminal groups interacting with the S3 and S1’ pockets respectively. B) Inhibition assay of Zika NS2B/NS3 protease by compound 86 gives an IC$_{50}$ of 0.958 μM, represented by the dotted line. Protease activity was measured by release of AMC. Assays were carried out in duplicate and measurements were taken in triplicate. Standard deviations are represented by error bars.
Figure 3 shows the results of the 50 ns molecular dynamics simulations and inhibition assay of compound 104. This compound had previously been tested against Dengue and West Nile Virus\textsuperscript{11}, but not yet against Zika. The simulations show that the arginine group in this compound disassociates from the S2 pocket, thus causing an unstable interaction with the protease binding pocket. The inhibition assay however shows that this compound had an IC\textsubscript{50} of 0.609 μM, which is lower than that of compound 86 and falls below 1 μM, indicating a potential lead for further development.

Figure 3. Molecular dynamics simulation and \textit{in vitro} enzymatic inhibition graph of compound 104 against Zika protease. A) 50 ns molecular dynamics simulation of compound 104 with Zika NS2B/NS3 protease. Green arrow indicates the shift of the arginine group out of the S2 pocket. B) Inhibition assay of Zika NS2B/NS3 protease by compound 104 gives an IC\textsubscript{50} of 0.609 μM, represented by the dotted line. Protease activity was measured by release of AMC. Assays were carried out in duplicate and measurements were taken in triplicate. Standard deviations are represented by error bars.

a) b)
After the previously discovered compounds had been tested and the enzyme assay was shown to be fit for purpose, additional novel compounds were examined to further understand the mechanism of action of these tripeptide inhibitors. Figure 4 shows the results for compound 9C, which has a very small cyclic C-terminal group indicated by the green arrow. The 50 ns molecular dynamics simulations show that this group disassociates from the S1’ pocket, as well as the arginine group disassociating from the S2 pocket, resulting in an overall unstable interaction. The IC$_{50}$ of this compound is 16.8 μM.

Figure 4. Molecular dynamics simulation and in vitro enzyme inhibition graph of compound 9C against Zika protease. A) 50 ns molecular dynamics simulation of compound 104 with Zika NS2B/NS3 protease. Green arrow indicates cyclic C-terminal group. B) Inhibition assay of Zika NS2B/NS3 protease by compound 104 gives an IC$_{50}$ of 16.8 μM, represented by the dotted line. Protease activity was measured by release of AMC. Assays were carried out in duplicate and measurements were taken in triplicate. Standard deviations are represented by error bars.
The next two compounds, 11C and 12C, are variants of 104 and 86 respectively. They contain a C-terminal histidine group. This was done in order to elucidate the binding pose of these molecules. A previous paper had reported that the C-terminal group interacted with the S1 pocket, shown in Figure 5a. However, previous simulations done in house (data not shown) indicated that this is an unstable conformation, and that the C-terminal group instead interacts with the S1’ pocket. A histidine C-terminal group was theorised to interact with the S1 pocket in the same way as previously described. However, Figure 5b shows that in the 50 ns molecular dynamics simulations this group (indicated by the arrow) also does not interact with the S1 pocket. During the purification of compound 11C, two different peaks were identified. Both these peaks were tested. The IC₅₀ values for compound 11C and 12C, shown in Figure 5c and Figure 6, are above 1 μM, indicating low inhibitory potential.

![Figure 5. Previously reported binding pose of 86 with Dengue and molecular dynamics simulation and in vitro enzyme inhibition graph of compound 11C against Zika protease. A) Previously reported docking pose of compound 86 from Behnam et al. C-terminal group indicated by arrow B) 50 ns molecular dynamics simulation of compound 11C with Zika NS2B/NS3 protease. Histidine group, indicated by arrow, shows no interaction with S1 or S1’ pocket. C) Inhibition assay of Zika NS2B/NS3 protease by compound 11C gives an IC₅₀ of 6.50 and 4.73 μM for peak 1 and 2 respectively, represented by the dotted line. Protease activity was measured by release of AMC. Assays were carried out in duplicate and measurements were taken in triplicate. Standard deviations are represented by error bars.](image-url)
The final compound to be tested was compound 21C (compound structure not shown due to confidentiality). This compound was a novel dipeptide inhibitor. This particular version of the dipeptide inhibitor was not optimised against the Zika NS2B/NS3 protease, as it has only one charged group (arginine) known to interact with the active site. Two docking poses (not shown) were produced. The first pose was calculated to have a lower free energy, and shows the C-terminal group interacting with the S2 pocket, the N-terminal group interacting with the S1 pocket and the arginine residue interacting with the S3 pocket. The second pose follows the trend seen with the tripeptide inhibitors, where the arginine residue interacts with the S2 pocket and the N-terminal group interacts with the S3 pocket. In the case of this compound however, the C-terminal group interacts with the S1 pocket. While this compound is not optimised against Zika, it still shows a sub-micromolar IC$_{50}$, indicating good inhibitory potential.

Figure 6. *In vitro* enzyme inhibition graph of compound 12C against Zika protease. Inhibition assay of Zika NS2B/NS3 protease by compound 12C gives an IC$_{50}$ of 13 \( \mu \text{M} \), represented by the dotted line. Protease activity was measured by release of AMC. Assays were carried out in duplicate and measurements were taken in triplicate. Standard deviations are represented by error bars.

![IC$_{50}$ = 13.0 \( \mu \text{M} \)](image)

Figure 7. *In vitro* enzyme inhibition graph of compound 21C against Zika protease. Inhibition assay of Zika NS2B/NS3 protease by compound 21C gives an IC$_{50}$ of 0.654 \( \mu \text{M} \), represented by the dotted line. Protease activity was measured by release of AMC. Assays were carried out in duplicate and measurements were taken in triplicate. Standard deviations are represented by error bars.

![IC$_{50}$ = 0.654 \( \mu \text{M} \)](image)
Discussion

The aim of this study was to explore the mechanism of action of recently discovered tripeptide Flavivirus NS2B/NS3 protease inhibitors using *in silico* molecular dynamics techniques and determine *in vitro* IC$_{50}$ values against Zika and TBE NS2B/NS3 proteases. Enzyme constructs were produced at the Protein Science Facility, Karolinska Institutet and the first step of this study aimed to set up an enzymatic assay to validate our protease constructs. Using a fluorescence based assay, we found that the unlinked TBE construct that was produced using a novel production method showed no activity. A mass spectrometry analysis (data not shown) of the construct revealed it was of the correct weight and it is therefore theorised that the TBE NS2B/NS3 protein did not fold correctly during expression. This taking account that the mass spectrometry analysis also showed that there was a high amount of bacterial proteins known to bind to the metal affinity chromatography and thereby contaminated the purification. Unlike the TBE construct, the Zika protease that had been produced following a previously published method$^{12}$ showed activity. Our enzyme kinetics assay showed that the Km of the Zika protease (Fig 1) was the same as to what was previously reported.$^{14}$

In addition to our fluorescence based assay, which could cause problems with fluorescent inhibitors, we attempted to use a colorimetric assay to measure the protease activity. However, the kit used (Pierce Colorimetric Protease Assay Kit, Thermo Scientific) did not work properly. This may have been the result of the provided substrate, succinylated casein, not being recognised by the protease, because NS2B/NS3 proteases have a specific consensus sequence preference.$^{10}$ In order to determine the effect of fluorescing inhibitors containing a bithiophene group on our assays, we examined them in combination with our reporter molecule, aminomethylcoumarin (AMC). We found that there was a linear relationship between the concentration of the inhibitor and the signal produced (data not shown) and therefore determined that our assay was fit for the purpose when the increased signal produced by the fluorescing inhibitor was accounted for.

Following the validation of the Zika protease, we investigated several inhibitors to determine IC$_{50}$ values. Figure 2b and Figure 3b show the IC$_{50}$ values of two previously published compounds, 86 and 104, which had shown good activity against Dengue and West Nile Virus (WNV) NS2B/NS3 protease$^{11}$. Compound 86 had been previously tested against Zika$^{13}$ and the IC$_{50}$ obtained in this study is the same as was previously reported. Compound 104 had not previously been tested against Zika, and while the IC$_{50}$ is higher than that reported for Dengue and WNV, it still falls in the sub-micromolar range. This makes it an interesting hit for further development. It is interesting to note that the 50 ns molecular dynamics (MD) simulations of compound 104 showed that the arginine residue disassociates from the S2 pocket (Fig. 3a), which does not happen with compound 86 (Fig 2a). As compound 104 has a lower IC$_{50}$ value, this might mean that this arginine residue may not be as important for inhibiting Zika Virus NS2B/NS3. Taken together, these results support the possibility that a pan-Flavivirus protease inhibitor could be developed, as presented in previous work by our group.$^9$ Together with the enzyme kinetics results, it also shows that our assays are fit for purpose and can be used for the investigation of novel compounds.

In addition to testing previously reported compounds against our Zika protease construct, we tested several novel compounds. The first series of compounds tested were based of the tripeptide scaffold that was previously published.$^{11}$ Figure 4 shows the MD simulation and inhibition graph of compound 9C, in which the C-terminal group has been replaced with a small cyclic group. Previously published data had suggested that reducing the
size of the C-terminal group has a negative effect on the inhibitory capabilities of this scaffold\textsuperscript{11}. The MD simulations indicated unstable binding of this compound and it was found that 9C also had a very high IC\textsubscript{50}. This supports the trend observed, indicating that the larger 4-Benzylxyloxy Phenyl GlycineC-terminal group may be important for the inhibitory capabilities of this scaffold.

Figure 5 and 6 shows the inhibition graph of compound 11C and 12C. These compounds contain a histidine C-terminal group with either a biothiophene or thiophene N-terminal group. They were tested to help elucidate the binding pose of the scaffold. In the original publication from which these molecules are derived, it was suggested that the C-terminal group interacts with the S1 pocket of the protease (Fig. 5a). However, previous work by our group showed that MD simulations using this pose were unstable. If the binding pose of the original publication is correct, replacing the C-terminal group with a histidine would allow it to interact with the S1 pocket in the same manner as the lysine group. The MD simulations of compound 11C (Fig. 5b) showed that the histidine group does not interact with the S1 or S1′ pockets and the \textit{in vitro} results (Fig. 5c and Fig. 6) show an increased IC\textsubscript{50} for both compounds. This supports the use of a binding pose in which the C-terminal interacts with the S1′ pocket.

Lastly, a novel dipeptide scaffold was tested against the Zika NS2B/NS3 protease. Figure 7a shows two binding poses for compound 21C. The first pose was calculated to have a lower free energy but the second pose follows the trend seen when using the tripeptide scaffold where the arginine residue interacts with the S2 pocket and the N-terminal group interacts with the S3 pocket. In this case the C-terminal group does not interact with the S1′ pocket but with the S1 pocket. Further MD simulations will be necessary to determine the stability of these two poses. Figure 7b shows the inhibition assay results of compound 21C, showing a sub-micromolar inhibitory concentration against Zika Virus NS2B/NS3. The structure of 21C is not optimised against Zika Virus NS2B/NS3 but with a sub-micromolar IC\textsubscript{50} it provides an interesting avenue for further development. Further variants of this scaffold have already been designed and will be investigated in the future.

While this is only a preliminary step in the development of a \textit{Flavivirus} antiviral, it is important to take into account the possibility of resistance development. While \textit{Flavivirus} proteases are highly conserved between species, and with the catalytic site recognizing very similar chemical environments, resistance could occur. This has been seen in Hepatitis C treatment, where certain mutations can even lead to cross-resistance to different types of protease inhibitors.\textsuperscript{16} However, once an effective drug has been developed, it will become easier to develop alternatives that could be used in a combination therapy to lower the chances of resistance occurring. Besides resistance, it is also important to think of the cost of treatment. As \textit{Flavivirus} infections mainly occur in developing countries, there will be issues with drug funding. However, as the burden of \textit{Flavivirus} infections is large, the investment in a drug would pay off. Furthermore, additional foreign funding could allow for the production of low cost drugs.

In conclusion, the two compounds 104 and 21C have been found to have sub-micromolar inhibitory concentrations. The tripeptide compound 104, which had been previously tested against Dengue and West Nile Virus, supports the idea of a pan-\textit{Flavivirus} NS2B/NS3 protease inhibitor while the novel dipeptide 21C provides a new scaffold for potential inhibitors. The dipeptide scaffold also has the advantage over the tripeptide scaffold of being more ‘drug-like’, which brings a potential Flavivirus antiviral therapy one step
closer. The binding mechanism of the tripeptide scaffold has also been investigated, providing useful information for the development of future novel inhibitors.

**Acknowledgments**

I would like to thank Johan Lennerstrand and Dario Akaberi for guiding me during the project and the Protein Science Facility at Karolinska Institute for providing the enzyme constructs. I would also like to thank Amanda Båhlström and Anja Sandström at the Department of Medicinal Chemistry for providing the inhibitors. Finally I would also like to thank the Zoonosis Science Center lab, IMBIM, and Uppmax, Uppsala University, for providing additional resources.

**References**

13. Kuiper, B. D. *et al.* Increased activity of unlinked Zika virus NS2B/NS3 protease

