Antiviral Research 139 (2017) 88–94

Contents lists available at ScienceDirect

Antiviral Research

journal homepage: www.elsevier.com/locate/antiviral

Extended substrate specificity and first potent irreversible inhibitor/activity-based probe design for Zika virus NS2B-NS3 protease

Wioletta Rut [a], Linlin Zhang [b, c], Paulina Kasperkiewicz [a, d], Marcin Poreba [a, d], Rolf Hilgenfeld [b, c], Marcin Drąg [a, *]

[a] Department of Bioorganic Chemistry, Faculty of Chemistry, Wroclaw University of Technology, Wyb. Wyspianskiego 27, 50-370 Wroclaw, Poland
[b] Institute of Biochemistry, Center for Structural and Cell Biology in Medicine, University of Lübeck, Ratzeburger Allee 160, 23562 Lübeck, Germany
[c] German Center for Infection Research (DZIF), Hamburg-Lübeck-Borstel-Riems Site, University of Lübeck, 23562 Lübeck, Germany
[d] NCI-designated Cancer Center, Sanford-Burnham Prebys Medical Discovery Institute, La Jolla, CA, 92037, USA

ARTICLE INFO

Article history:
Received 22 November 2016
Received in revised form
19 December 2016
Accepted 21 December 2016
Available online 26 December 2016

Keywords:
Irreversible inhibitor
Proteolytic enzyme
ZIKA
Substrate specificity

ABSTRACT

Zika virus is spread by *Aedes* mosquitoes and is linked to acute neurological disorders, especially to microcephaly in newborn children and Guillan-Barré Syndrome. The NS2B-NS3 protease of this virus is responsible for polyprotein processing and therefore considered an attractive drug target. In this study, we have used the Hybrid Combinatorial Substrate Library (HyCoSuL) approach to determine the substrate specificity of ZIKV NS2B-NS3 protease in the P4-P1 positions using natural and a large spectrum of unnatural amino acids. Obtained data demonstrate a high level of specificity of the S3-S1 subsites, especially for basic amino acids. However, the S4 site exhibits a very broad preference toward natural and unnatural amino acids with selected *D*-amino acids being favored over *L* enantiomers. This information was used for the design of a very potent phosphonate inhibitor/activity-based probe of ZIKV NS2B-NS3 protease.

1. Introduction

Zika virus (ZIKV) was first isolated from rhesus monkey cells and initially characterized by Dick in 1947 (Dick et al., 1952). Shortly thereafter, this virus spread in tropical and subtropical regions of central Africa as well as South and Southeast Asia. Recently, ZIKV has been rapidly spreading in the Americas and the Carribean. Recent studies have revealed that Zika virus is responsible not only for mild and self-limiting illnesses as rash, fever, myalgia but can also cause serious neurological disorders such as Guillain-Barré syndrome (Fontanet et al., 2016) and microcephaly in newborn children (Rasmussen et al., 2016). There is currently no vaccine or antiviral drug to prevent or treat ZIKV infection.

ZIKV belongs to the *Flaviviridae* family of RNA viruses. Its 10.7-kb genome encodes a single polyprotein, which is split into three structural proteins (C, prM/M, E) and seven nonstructural proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B, NS5) (Lei et al., 2016). The N-terminal third of the NS3 protein is a chymotrypsin-like serine protease that, in collaboration with the NS2B protein, cleaves the viral polyprotein C-terminal to dibasic sites, and is thus required for viral replication. Association of NS3 with the membrane-bound NS2B protein is necessary to achieve proteolytic activity, similar to what has been observed for the related Dengue virus protease (Leung et al., 2001; Lindenbach and Rice, 2003). Thus, the NS2B/NS3 protease (NS2B/NS3^{pro}) appears as a very attractive molecular target for the treatment of ZIKV infection. Recently, the crystal structure of ZIKV NS3^{pro} along with a minimum version of NS2B (i.e. the hydrophilic region of residues 49* - 95*; the asterisk denotes residues of NS2B) and a boronate inhibitor has been published (Lei et al., 2016).

In the present study, we used the recently developed Hybrid Combinatorial Substrate Library (HyCoSuL) approach to identify the extended substrate specificity of ZIKV protease in the S4-S1 substrate-binding subsites. The use of natural and a large pool of unnatural amino acids in the substrate library allowed us to fully explore the chemical space compatible with the NS2B-NS3^{pro} catalytic cleft, enabling us to design and synthesize a first-in-class fluorogenic substrate and a very potent phosphonate inhibitor/activity-based probe.

* Corresponding author.
E-mail address: marcin.drag@pwr.edu.pl (M. Drąg).

http://dx.doi.org/10.1016/j.antiviral.2016.12.018
0166-3542/

2. Materials and methods

Synthesis of all compounds and their structural analysis, P1 library synthesis, substrates and activity-based probe kinetic analysis can be found in Supporting Information.

2.1. Visualization of ZIKV NS2B–NS3 protease with the WRPK3 activity-based probe

The active-site concentration was measured by titration of the enzyme with the phosphonate WRPK3 probe. Enzyme was incubated with a broad range of probe (500 - 4 nM) for 30 min ($t_{1/2}$<3s) followed by substrate addition. RFU values were measured in time (SpectraMax Gemini XPS, Molecular Devices) using an excitation wavelength of 355 nm and an emission wavelength of 460 nm (cutoff 455 nm). Next, the enzyme concentration was calculated from the dependence of RFU/s from the inhibitor concentration. Afterwards, the probe was diluted to a final concentration in the range 500–4 nM. The first dilution was made in DMSO and then in NS2B–$NS3^{pro}$ assay buffer (10 mM Tris-HCl, 20% glycerol, 1 mM CHAPS, 1 mM DTT, pH = 8.5). Final DMSO concentration was less than 2%. The enzyme was diluted to a final concentration of 100 nM in NS2B–$NS3^{pro}$ assay buffer. Aliquots of 50 μL of enzyme were incubated with 50 μL of various concentrations of probe (or buffer as the control) for 10 min at room temperature, followed by denaturation with 50 μL of 3xSDS loading buffer with DTT for 5 min at 95°. As a standard, SDS-PAGE Molecular Weight Standards, Broad Range BIO-RAD was used. SDS-PAGE was performed using Bolt 4%–12% Bis-Tris Plus Gels (Novex, Life Technologies) for 30 min with constant voltage (200 V) at room temperature, followed by transfer to the nitrocellulose membrane (0.2 μM; Bio-Rad) for 60 min (10 V). The membrane was blocked for 1 h with 2% BSA in TBS-T buffer (Tris-buffered saline with Tween 20, v/v) at room temperature. Afterwards, to visualize the biotinylated probe, the membrane was treated with IRDye 800 CW Streptavidin (1:10 000) for 30 min at room temperature and visualized with an Odyssey fluorescence imaging system (LI-COR).

3. Results

3.1. Enzyme preparation

The ZIKV NS2B-NS3 protease preparation method used in this study was almost identical to that described by Lei et al. (2016). The hydrophilic region, residues 49* - 95*, of the cofactor NS2B was covalently linked via a $Gly_4SerGly_4$ sequence to the N-terminus of NS3, which comprised residues 1–170. This kind of construct had originally been introduced for Dengue virus (DENV) protease by Leung et al. (2001) and was found to be highly active for ZIKV protease (Lei et al., 2016). Two mutations, Arg95*Ala and Arg29Gly were introduced to reduce possible autolysis of the protease.

3.2. Substrate specificity

To investigate the NS2B-NS3 protease substrate specificity at the S4-S2 subsites, we used a previously synthesized HyCoSuL P1-Arg library (Kasperkiewicz et al., 2015). This library contains 19 natural (omitting cysteine) and 108 unnatural amino acids in each position of three sub-libraries (P4, P3, and P2) (details on synthesis in Supporting Information). In order to determine the P1 specificity, we synthesized a library of general structure Ac-Ala-Arg-Leu-X-ACC, with constant P4-P2 sequence and different natural and unnatural amino acids at the P1 position (details on synthesis in Supporting Information) using approach described by us earlier (Kasperkiewicz et al., 2014).

3.3. P1 specificity

The S1 pocket of NS2B-NS3 protease has very narrow specificity. The most preferred amino acids are L-Arg(Me) and L-Arg (Fig. 1). These two basic amino acids are recognized almost at the same level. L-Lys is also recognized, however its activity is around three times lower. Other amino acids including natural and unnatural demonstrate little to no activity. These data clearly demonstrate that the S1 pocket preferentially recognizes long, basic amino acids.

3.4. P2 specificity

In this position, the NS2B-NS3 protease displays also a very high level of specificity (Fig. 2). The most preferred amino acids are the non-proteinogenic L-Orn (100%) and L-$Arg(Z)_2$ (40%). Natural amino acids interact significantly weaker with the S2 pocket, and only L-Lys and L-Arg are recognized (38% and 17%, respectively). Interestingly, both L-Arg and its homolog elongated by one methylene group, L-hArg, are recognized at the same level. Similarly to the S1 pocket, S2 also accepts only long, basic amino acids.

3.5. P3 specificity

Similarly to the S2 and S1 pockets, the S3 site exhibits very narrow specificity (Fig. 2). L-Lys is the most preferentially recognized among all tested amino acids. Other amino acids such as L-Orn, L-Arg, L-$Arg(Z)_2$, L-hArg, L-Agp or L-Phe(guan) are recognized with significantly lower affinity. However, all these amino acids have basic side chains. Residual activity (lower than 2%) of several other amino acids is also observed.

3.6. P4 specificity

Analysis of this position demonstrates that the S4 site exhibits very broad substrate specificity (Fig. 2). Almost all amino acids, including unnatural ones, are recognized at the same level. Surprisingly, the two most preferred amino acids are the unnatural D-Arg and D-Lys. Also, other D-amino acids are fairly well tolerated in the P4 position. Interestingly, only few amino acids (mostly negatively charged ones) are completely rejected in this position by the ZIKV NS2B-NS3 protease.

3.7. Kinetic constants for ZIKV NS2B-NS3 protease substrate cleavage

In the next step, we used the P4-P1 substrate specificity data for validation of our combinatorial approach as well as for the design of an optimal substrate sequence. Thus, we selected sequences with optimal natural and unnatural amino acids in each position (WR3, WR4, and WR5), a sequence with only natural amino acids (WR1) and a sequence used in previous studies (WR2) (Table 1) (Lei et al., 2016). Each sequence was C-terminally labeled with the ACC fluorophore to give a substrate, which was further purified and analyzed by HPLC (details on synthesis in Supporting Information). Next, kinetic parameters (k_{cat}, K_m, k_{cat}/K_m) for the synthesized substrates were measured. The data obtained very well matched the HyCoSuL substrate specificity profiling. The best substrates were those with D-amino acids in the P4 position. The highest k_{cat}/K_m value was displayed by Ac-*D*Arg-Lys-Orn-Arg-ACC (WR3), but Ac-*D*Lys-Lys-Orn-Arg-ACC (WR4) and Ac-*D*Arg-Lys-Orn-Arg(Me)-ACC (WR5) were only slightly less active (Table 1). The comparison of the P1 L-Arg substrate (WR3) with its mono-methylated analogue (Arg(Me)) (WR5) also mirrors the HyCoSuL screening data. WR3-WR5 substrates were also around three times more active compared to the previously reported sequence Ac-Nle-Lys-