Yellow fever virus (YFV) NS3 protease: peptide-inhibition studies


Novartis Institute for Tropical Diseases, 10 Biopolis Road, #05-01 Chromos, Singapore 138670

A recombinant form of yellow fever virus (YFV) NS3 protease, linked via a nonapeptide to the minimal NS2B co-factor sequence (CF40-gly-NS3pro190), was expressed in *Escherichia coli* and shown to be catalytically active. It efficiently cleaved the fluorogenic tetrapeptide substrate Bz-norleucine-lysine-arginine-arginine-AMC, which was previously optimized for dengue virus NS2B/3 protease. A series of small peptidic inhibitors based on this substrate sequence readily inhibited its enzymic activity. To understand the structure–activity relationship of the inhibitors, they were docked into a homology model of the YFV NS2B/NS3 protease structure. The results revealed that the P1 and P2 positions are most important for inhibitor binding, whilst the P3 and P4 positions have much less effect. These findings indicate that the characteristics of YFV protease are very similar to those reported for dengue and West Nile virus proteases, and suggest that pan-flavivirus NS3 protease drugs may be developed for flaviviral diseases.

INTRODUCTION

Yellow fever virus (YFV) is a mosquito-borne flavivirus with a present-day distribution confined largely to equatorial Africa and central South America. The enveloped virus contains a single- and positive-stranded RNA genome that encodes a polyprotein of over 350 kDa (Rice *et al.*, 1985; Chambers *et al.*, 1990a) with the following gene order: 5′- C-prM-E-NS1-NS2A-NS2B-NS3-NS4A-NS4B-NS5-3′. C, prM and E are structural proteins and NS1–NS5 are non-structural proteins involved in polyprotein processing and replication. Numerous studies have shown that the YFV NS3 protease, in complex with NS2B, is involved in co- and post-translational processing of the viral polyprotein and recognizes cleavage sequences (G/ARR ↓ S/G) with consensus dibasic residues at the P1 and P2 positions to generate the N termini of NS2B, NS3, NS4A and NS5, and additional cleavage sites within core, NS2A and NS4A (Chambers *et al.*, 1991, 1993, 1995, 2005; Amberg *et al.*, 1994; Droll *et al.*, 2000).

The YFV NS3 protease belongs to the trypsin superfamily and is relatively well conserved amongst members of the family *Flaviviridae* (Bazan & Fletterick, 1989; Gorbalenya *et al.*, 1989). The essentiality of NS3 protease activity in viral replication, through at least its polyprotein-cleaving activity, has been demonstrated by mutational analysis of the residues predicted to form the canonical serine protease catalytic triad (H53, D77 and S138) (Chambers *et al.*, 1990b, 1993, 1995, 2005; Droll *et al.*, 2000).

An effective YFV vaccine was developed when the disease caused havoc in North America at the dawn of the 20th century. There still remains a chance that the disease can re-emerge in outbreak proportion because of its increased incidence in the past 25 years and risks of urban YFV in Africa and South America (Barrett & Higgs, 2007). This has led to calls for the development of chemotherapeutic agents for yellow fever so as not to be caught unprepared. Presently, there is no chemotherapy for any of the flaviviral infections.

The NS3 protease, being highly conserved amongst flaviviruses, would be an attractive potential target for a pan-flaviviral drug. Considerable efforts have been made recently to find such inhibitors to both dengue virus (DENV; Leung *et al.*, 2001; Ganesh *et al.*, 2005; Li *et al.*, 2005; Yin *et al.*, 2006a, b) and West Nile virus (WNV; Nall *et al.*, 2004; Ganesh *et al.*, 2005; Erbel *et al.*, 2006; Knox *et al.*, 2006) proteases. Similar undertakings have not been reported for the YFV protease. In this context, understanding the molecular interaction between flaviviral NS3 protease and inhibitors is a question of broad relevance and may facilitate the development of designed drugs. As DENV, WNV and YFV NS2B/NS3 proteases share similar cleavage recognition sites (Chambers *et al.*, 1990a; Falgout *et al.*, 1991), we reasoned that peptidic inhibitors generated against DENV (Yin *et al.*, 2006a, b) and WNV (Erbel *et al.*, 2006).
proteases could potentially act on the YFV protease. To test this hypothesis, we studied the effects of these inhibitors on YFV NS3 protease activity by using a single-chain complex of YFV NS2B/NS3 protease (CF40-gly-NS3pro190) comprising 47 core amino acids of NS2B (Chambers et al., 1991; Falgout et al., 1993) linked via Gly4–Ser–Gly4 to the N-terminal 190 aa NS3 protease domain. Similar designs in DENV (Leung et al., 2001), WNV (Nall et al., 2004) and YFV (Bessaud et al., 2006) produced catalytically active proteases.

YFV CF40-gly-NS3pro190 was expressed as an N-terminally His-tagged fusion protein in *Escherichia coli* and purified from an Ni²⁺ affinity column from clarified bacterial lysate [Supplementary Figure S1(a), available in JGV Online]. The total amount of active protein in the purified fraction was 100%, as determined after active-site titration with the competitive, tight-binding inhibitor aprotinin (*K*ₐₕ = 17.61 ± 0.56 nM; Supplementary Figure S2, available in JGV Online) and curve fitting according to the Morrison equation (Copeland, 2000; Supplementary Figure S2). The proteolytic activity of YFV CF40-gly-NS3pro190 was first assessed with two different fluorogenic substrates: the tripeptide Boc-GRR-AMC (Yusof et al., 2000; Bessaud et al., 2006) and the tetrapeptide Bz-nKRR-AMC (Li et al., 2005). We observed that its catalytic efficiency for Boc-GRR-AMC (*k*ₐₕ/*K*ₐₕ = 239.05 ± 7.7 M⁻¹ s⁻¹; Table 1) was comparable to that reported by Bessaud et al. (2006) (*k*ₐₕ/*K*ₐₕ = 105 M⁻¹ s⁻¹). However, the enzyme exhibited an almost 10-fold higher affinity for the tetrapeptide substrate and also processed it with 34-fold greater efficiency (Table 1). This finding may in part be due to the additional P4 residue (norleucine) in the tetrapeptide sequence. In DENV, tetrapeptides were also shown to give improved substrate binding (giving rise to a better *K*ₐₕ value) and turnover (better *k*ₐₕ value) compared with tripeptides (Niyomrattanakit et al., 2006). The tetrapeptide substrate was thus employed for subsequent enzyme characterization and inhibitory studies with peptidic inhibitors. No activity was obtained with an S138A mutant construct of YFV CF40-gly-NS3pro190 when tested against Bz-nKRR-AMC at enzyme concentrations ranging from 0 to 1000 nM and even after incubation with the substrate for 6 h (Supplementary Figure S3, available in JGV Online). This indicates that the enzymatic activity of wild-type YFV CF40-gly-NS3pro190 was not due to contamination by bacterial proteases.

We determined that the optimal buffer conditions for YFV protease activity were high pH (8.5), low ionic strength and the presence of divalent cations (Supplementary Figure S4, available in JGV Online). These buffer requirements resemble those reported for DENV (Leung et al., 2001), WNV (Nall et al., 2004) and YFV (Bessaud et al., 2006). Furthermore, similar to reports on DENV (Leung et al., 2001) and WNV (Nall et al., 2004), YFV activity was also enhanced by addition of glycerol (Supplementary Figure S4, available in JGV Online). Based on these results, we used these buffer conditions (50 mM Tris/HCl, pH 8.5; 1 mM CHAPS; 20% glycerol) to assay the properties of the peptidic inhibitors.

We first employed the aldehyde derivative of the substrate-based peptide Bz-nKRR-H (1; Table 2), which has *K*ᵢ values of 5.8 and 2.1 μM for DENV (Yin et al., 2006a, b) and WNV (Erbel et al., 2006; Knox et al., 2006), respectively.

### Table 1. Kinetic parameters for the hydrolysis of substrates by YFV CF40-gly-NS3pro190

Steady-state kinetic parameters (*K*ᵢ, *k*ₐₕ and *k*ₐₕ/*K*ᵢ) of YFV CF40-gly-NS3pro190 were obtained with Boc-Gly-Arg-Arg-AMC (*n* = 1) and Bz-Nle-Lys-Arg-Arg-AMC (*n* = 5). Hydrolysis was performed at 37 °C in 50 mM Tris/HCl (pH 8.5), 1 mM CHAPS and 20% glycerol with substrate concentration from 250 to 5 μM. Reaction progress was monitored continuously at 37 °C for 15 min by following the increase in fluorescence min⁻¹ (excitation, 385 nm; emission, 465 nm) on a Victor 3 plate reader (Perkin-Elmer). Triplicate measurements were taken for each data point and the data are reported as means ± SEM.

<table>
<thead>
<tr>
<th>Substrate</th>
<th><em>K</em>ᵢ (μM)</th>
<th><em>k</em>ₐₕ (s⁻¹)</th>
<th><em>k</em>ₐₕ/<em>K</em>ᵢ (s⁻¹ M⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boc-GRR-AMC</td>
<td>142.2 ± 9.2</td>
<td>0.034 ± 0.001</td>
<td>239.05 ± 7.7</td>
</tr>
<tr>
<td>Bz-nKRR-AMC</td>
<td>14.6 ± 1.7</td>
<td>0.111 ± 0.018</td>
<td>8024 ± 1537</td>
</tr>
</tbody>
</table>

Table 2. Inhibition of YFV CF40-gly-NS3pro190 by peptidic inhibitors

Peptidic inhibitors in serially diluted concentrations ranging from 0 to 100 μM were pre-incubated with YFV CF40-gly-NS3pro190 for 30 min before assaying with substrate Bz-nKRR-AMC in 50 mM Tris/HCl (pH 8.5), 1 mM CHAPS and 20% glycerol (Yin et al., 2006a, b; Knox et al., 2006). Alanine (2–5) and phenylalanine (6–9) scans in the P4–P1 positions, Lys replacement in the P1 and P2 positions (10–11), replacement with boronic acid warhead (14) and shortened inhibitors (12, 13) were carried out to assess the importance of the P4–P1 positions on the inhibitory properties of Bz-nKRR-H. Results are expressed as *K*ᵢ values and are means of three or more independent experiments.

### Table 2. Inhibition of YFV CF40-gly-NS3pro190 by peptidic inhibitors

<table>
<thead>
<tr>
<th>Peptide-based inhibitor</th>
<th><em>K</em>ᵢ (μM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Bz-Nle-Lys-Arg-Arg-H</td>
<td>0.40 ± 0.09</td>
</tr>
<tr>
<td>2. Bz-Nle-Lys-Arg-Ala-H</td>
<td>33.15 ± 2.21</td>
</tr>
<tr>
<td>4. Bz-Nle-Ala-Arg-Arg-H</td>
<td>0.90 ± 0.10</td>
</tr>
<tr>
<td>5. Bz-Ala-Lys-Arg-Arg-H</td>
<td>0.75 ± 0.11</td>
</tr>
<tr>
<td>6. Bz-Nle-Lys-Arg-Phe-H</td>
<td>62.00 ± 3.00</td>
</tr>
<tr>
<td>7. Bz-Nle-Lys-Phe-Arg-H</td>
<td>15.60 ± 2.00</td>
</tr>
<tr>
<td>8. Bz-Phe-Arg-Arg-H</td>
<td>1.50 ± 0.31</td>
</tr>
<tr>
<td>9. Bz-Phe-Lys-Arg-Arg-H</td>
<td>0.25 ± 0.06</td>
</tr>
<tr>
<td>11. Bz-Nle-Lys-Lys-Arg-H</td>
<td>1.20 ± 0.10</td>
</tr>
<tr>
<td>12. Bz-Lys-Arg-Arg-H</td>
<td>0.40 ± 0.01</td>
</tr>
<tr>
<td>13. Bz-Arg-Arg-H</td>
<td>0.85 ± 0.06</td>
</tr>
<tr>
<td>14. Bz-Nle-Lys-Lys-Arg-(OH)₂</td>
<td>0.05 ± 0.01</td>
</tr>
</tbody>
</table>
Against YFV, it was almost 15-fold more potent than DENV and 5-fold better than WNV ($K_i = 0.4 \mu M$; Table 2). Its potency could be improved by 8-fold by substituting the aldehyde warhead with the more electrophilic group boronic acid (14; Table 2). Replacement of P1 (Arg) in Bz-nKRR-H with Ala (2; Table 2) reduced its potency dramatically by 83-fold. Replacement with Phe (6; Table 2) had an even more pronounced effect (155-fold decrease). Interestingly, the P1 position could not be substituted with positively charged Lys, as this reduced potency by 28-fold (10; Table 2). Substitution of P2 Arg with Ala resulted in a complete loss of inhibition (3; Table 2), whilst replacement with Phe caused a 39-fold decrease (7; Table 2). Unlike in the P1 site, substitution with Lys in the P2 site was tolerated, as enzyme activity was reduced marginally by about 3-fold (11; Table 2). Substitutions of P3-Lys and P4-Nle with either Ala or Phe did not affect inhibitory activity significantly (4, 5, 8, 9; Table 2). Likewise, when the peptide-inhibitor sequence was shortened to tri- and dipeptides, $K_i$ values also did not change significantly (12, 13; Table 2). Taken together, these data show that S1 and S2 are the major determinants for the binding of the peptidic inhibitor.

To understand the observed inhibitory properties of these peptidic inhibitors against the YFV protease, we built a homology model of the YFV NS2B/NS3 protease (Fig. 1) based on the crystal structure of WNV NS2B/NS3 in complex with Bz-nKRR-H (2FP7; Erbel et al., 2006) and replaced the latter with individual peptidic aldehydes by using SYBYL 6.9 followed by optimization with MacroModel 9.0, to study their effects on interactions with amino acids in the YFV active site (Supplementary Methods, available in JGV Online).

Homology modelling was performed by using Prime 1.2 (Schrödinger). The model showed that the S1 and S2

---

**Fig. 1.** (a) Structural alignment of WNV and YFV NS2B/NS3 sequences, indicating amino acids conserved in both proteases (blue text), S1–S4 active-site pockets (yellow boxes) and the catalytic triad (red asterisks). Residues lacking electron density in the crystal structure of WNV NS2B/NS3 are shown in green and italicized. (b) Homology model of the active site of the YFV protease (see Supplementary Methods, available in JGV Online), showing Bz-nKRR-H (hydrogen removed for clarity) in the S1–S4 pockets, with hydrogen-bond donors (red) and hydrogen-bond acceptors (blue) painted on a Connolly surface and the catalytic triad (S138, H53 and D77) also indicated. (c) Putative interactions of YFV NS3 protease and Bz-nKRR-H with residues from YFV NS3 (in bold type) and NS2B (in italics).
pockets of the YFV protease were the most well defined compared with the S3 and S4 pockets (Fig. 1b). A large portion of the S1 pocket comprised the aromatic side chains of Y153 and F164 (Fig. 1c). As was observed in the WNV protease, a cation–π stacking occurred, with the guanidine group from P1-Arg sandwiched between the phenyl rings of the Bz cap and the inhibitor and F164. Two additional residues, D132 and G154, formed key interactions with P1-Arg via its side chain and backbone, respectively (Fig. 1c). The S2 pocket was narrower and shallower than S1 and was lined with hydrogen-bond donors and acceptors from NS2B and the NS3 domain, with three hydrogen bonds formed between S2 and P2-Arg: two with NS3 (D77 and N155) and one with NS2B (G82; Fig. 1c). The S3 pocket was larger, but less well defined, than S1 and S2, again comprising residues from both NS2B and the NS3 domain. Interestingly this pocket possessed a dual character in which half of the pocket was hydrophobic (F84 from NS2B; I157, L158 from NS3), whilst the other half comprised hydrogen-bond donor/acceptor atoms from the backbone of E83, K85 (NS2B) and G156 (NS3). P3-Lys also formed three hydrogen bonds; the first was a backbone hydrogen bond to G156 in NS3, whilst the remaining two were side-chain hydrogen bonds to E83 and F84 in NS2B (Fig. 1c). The S4 pocket was the least well defined and was made up entirely of L158, yielding a hydrophobic, solvent-exposed patch on the surface (Fig. 1c). No hydrogen bonds were found between the side chain of L158 and P4-Nle. Overall, the YFV homology model resembled the WNV structure closely, in that the S1 and S2 pockets were the dominant recognition sites for the peptidic inhibitor (Erbel et al., 2006).

We next replaced Bz-nKRR-H in the YFV homology model with the other peptidic inhibitors tested. We observed that side-chain interactions with charged and aromatic residues lining the S1 and S2 pockets were lost when P1 and P2 Arg were replaced with Ala or Phe in peptidic inhibitors 2, 3, 6 and 7, leading to dramatic drops in the inhibitory activity (>40-fold, Table 2; D132 for 2 and 6; G82 and N155, D77 for 3 and 7). Furthermore, these substitutions disrupted the π-stacking of P1-Arg, the Bz cap and F164. Although P1-Ala substitution (2) resulted in loss of the same hydrogen bonds as P1-Phe (6), there was a nearly 2-fold difference in their inhibitory activities (Table 2). One explanation for this difference may be that the smaller size of Ala allowed the hydrophobic Bz cap to bury more deeply into the S1 pocket, giving rise to a more stable complex than Bz-nKRR-H.

In contrast, the inhibitory activity of Bz-nKKR-H (11) was comparable to that of 1. Bz-nKKR-H retained the key hydrogen bonding with D77, and its P2-Lys terminal NH$_3^+$ interacted similarly with G82 (NS2B domain) as P2-Arg in Bz-nKRR-H (1). In the case of Bz-nKKR-H (10), the P1-Lys NH$_3^+$ group failed to form the optimal cation–π interaction found in Bz-nKRR-H (1). Nevertheless, it retained the hydrogen bond with the side chain of D132; this probably explains why it is less potent than 1, but more potent than the corresponding Ala (2) and Phe (6) substitutions in P1 (Table 2). The WNV protease also shows the same Lys tolerance in P2, but not in P1 (Erbel et al., 2006; Knox et al., 2006). However, for the DENV protease, Lys substitutions in both P1 and P2 are not tolerated (Yin et al., 2006b), suggesting that there may be differences in the DENV S1 pocket. This was supported by recent findings reported by Chappell et al. (2006), showing that the WNV protease was twice as active against substrates with Lys rather than Arg in P2, whilst DENV protease behaved in the opposite manner.

Side-chain interactions of S3 residues E83 and F84 with P3-Lys were lost with Ala or Phe substitutions (4, 8), whilst backbone hydrogen-bonding interaction with G156 was retained. As peptidic inhibitors 4 and 8 and the tripeptidic inhibitor (12) did not exhibit reduced potency, these side-chain interactions are probably not important. Finally, studies with both the dipeptide inhibitor and P4 replacements with Ala and Phe (5, 9) also showed that norleucine did not play a significant role in inhibitor binding and are consistent with the homology model, where the P4 residue does not form any hydrogen bonds with S4 residues.

In summary, the YFV protease shares similar characteristics with the DENV and WNV proteases and can be inhibited by the same peptidic inhibitor, Bz-nKRR-H. In all three cases, inhibitor potency is driven predominantly by interactions with amino acid residues in the S1 and S2 pockets, as substitutions in the P1 and P2 sites have the greatest impact on inhibitor potency. The information from these studies suggests that the design of pan-flavivirus drugs against NS2B/3 protease may be possible and that the peptidic inhibitors can serve as useful starting points for rational drug design.

Acknowledgements

We would like to thank Dr Charles Rice, Rockefeller University, New York, USA, for the kind gift of plasmid pYF M5-2, and David Beer and Viral Patel for advice and helpful discussions.

References


Chambers, T. J., Weir, R. C., Grakoui, A., McCourt, D. W., Bazan, J. F., Fletterick, R. J. & Rice, C. M. (1990b). Evidence that the N-terminal...
domain of nonstructural protein NS3 from yellow fever virus is a serine protease responsible for site-specific cleavages in the viral polyprotein. Proc Natl Acad Sci U S A 87, 8898–8902.


