Peptide Inhibitors of West Nile NS3 Protease: SAR Study of Tetrapeptide Aldehyde Inhibitors

John E. Knox, Ngai Ling Ma, Zheng Yin, Sejal J. Patel, Wei-Ling Wang, Wai-Ling Chan, K. R. Ranga Rao, Gang Wang, Xinyi Ngew, Viral Patel, David Beer, Siew Pheng Lim, Subhash G. Vasudevan, and Thomas H. Keller*

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

Received June 26, 2006

A series of inhibitors related to the benzoyl-norleucine-lysine-arginine-arginine (Bz-nKRR) tetrapeptide aldehyde was synthesized. When evaluated against the West Nile virus (WNV) NS3 protease, the measured IC_{50} ranges from ~1 to 200 μM. Concurrently, a modeling study using the recently published crystal structure of the West Nile NS3/NS2B protease complex (pdb code 2FP7) was conducted. We found that the crystal structure is relevant in explaining the observed SAR for this series of tetrapeptides, with the S1 and S2 pockets being the key peptide recognition sites. In general, a residue capable of both π-stacking and hydrogen bonding is favored in the S1 pocket, while a positively charged residue is preferred in the S2 pocket. This study not only confirms the importance of the NS2B domain in substrate-based inhibitor binding of WNV, it also suggests that the crystal structure would provide useful guidance in the drug discovery process of related *Flavivirus* proteases, given the high degree of homology.

Introduction

West Nile virus (WNV) was first identified in 1937 in the West Nile district of Uganda and was subsequently characterized in the 1950s.[1] Since then, the virus has spread to the different continents, with outbreaks in Algeria, Romania, Russia, etc.[2] The virus was first found in North America in 1999 and the subsequent spread in the United States is an important milestone in the evolving history of this virus.[3] Like many other viruses in the *Flavivirdae* family, WNV is transmitted by the bite from infected mosquitoes.[4,5] While 80% of the infected human population shows no symptoms, the remaining population may develop mild symptoms and in severe cases the disease may be fatal.[6]

WNV contains single-stranded, positive-sense RNA genome, which encodes three structural (C, prM, E) and seven nonstructural (NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5) proteins.[7] At present, there is no specific treatment for WNV infection. An attractive target identified for drug development is the WNV NS3 protease. Apart from being essential for viral replication, success has been demonstrated with the closely related hepatitis C virus (HCV) protease.[8−11] For flavivirus NS3 protease activity, it has been demonstrated that the NS2B cofactor is essential for catalytic activity.[12,13] With the lack of crystal structure, rationale-based drug design of WNV protease inhibitors has been guided by homology models: the NS3 domain either alone[14,15] or together with the NS2B domain, based on the crystal structure of HCV NS3 with NS4A cofactor.[16,17] However, the protease activity of HCV only requires a short 14 amino acid segment of the NS4A, as opposed to a 40 amino acid segment in other flaviviruses like WNV.[18,19] This throws into question whether these homology models are appropriate representations of the active WNV protease.

However, very recently a high-quality crystal structure of the WNV NS2B/NS3 cocrystallized with the substrate-based inhibitor benzoyl-norleucine-lysine-arginine-arginine (Bz-nKRR) aldehyde has become available from our group.[20] This structure revealed that the WNV NS2B cofactor differs substantially from other cofactor activated proteases such as HCV NS4A/NS3 and is actively involved in substrate binding. Motivated by the availability of this interesting structure, we conducted a modeling study in order to better understand the interactions between the WNV NS2B/NS3 and peptide-based inhibitors. By comparing the models with the in vitro activities of these inhibitors, we found that the crystal structure is relevant in explaining the observed SAR for this series of tetrapeptide inhibitors. This not only confirms the importance of the NS2B domain in substrate-based inhibitor binding of WNV, it also suggests that the crystal structure would provide useful guidance in the drug discovery process for related *Flavivirus* proteases.

Methods

To probe the binding characteristics of the WNV protease active site, a series of tetrapeptide aldehydes was synthesized (see our previously published work for synthetic details)[21,22] and tested for activity. Inhibitors were assayed in a 96-well plate format using 50 mM Tris, pH 8.5, 1 mM CHAPS in a final volume of 50 μL. Typically enzyme conjugate WNV CF40-gly-NS3pro187 (50 nM) was preincubated with various concentrations of test compounds at 37 °C for 30 min. The reaction was then initiated by the addition of substrate Bz-nKRR-AMC at 20 μM. Reaction progress was monitored continuously by following the increase in fluorescence (excitation 385 nm, emission 465 nm) on a TECAN Safire plate reader. IC_{50} values were derived for inhibitors by fitting the calculated initial velocities to a nonlinear regression curve fit using GraphPad Prism software (San Diego, CA). Each point of the IC_{50} curve was carried out in duplicate during a single experiment.

The synthesized peptide aldehydes (Table 1) were modeled using the crystal structure of WNV NS2B−NS3pro cocrystallized with the tetrapeptide aldehyde Bz-nKRR-H (PDB ID 2FP7) that was recently published by our group.[20] Preparation of the protein prior to modeling included adding hydrogen via the "Add Hydrogen" utility in Sybyl 6.9.[23] Without perturbing the protein structure, the Bz-nKRR tetrapeptide template was mutated to the corresponding peptide aldehydes in Table 1 using Sybyl 6.9.[23] These mutated inhibitor−protein complexes were then imported into Maestro 7.0 for optimization.[24] The prepared geometries corresponded to the protease tetrahedral transition structure with the aldehyde oxygen of the inhibitor given a formal charge of −1 and assumed to be singly bonded to the carbonyl carbon. For the catalytic triad S135-H51-D75 of the WNV protease, the charge on H51 is assumed to be positive with the addition of the proton from S135 to the imidazole ring while the D75 side chain is negatively charged. All

* Corresponding author. Phone: 65 6722 2908. Fax: 65 6722 2910. E-mail: thomas.keller@novartis.com.

10.1021/jm0607606 CCC: $33.50
Published on Web 10/05/2006

Table 1. Inhibition of West Nile Virus NS3 Protease by Tetrapeptide Aldehyde Inhibitors[a]

compd	aldehyde inhibitors	IC_{50} (mM)
1	Bz-Nle-Lys-Arg-Arg-H	4.1
2	Bz-Nle-**Pro**-Arg-Arg-H	14.3
3	Bz-Nle-Lys-**N-Me-Arg**-Arg-H	57.8
4	Bz-Nle-**N-Me**-Lys-Arg-Arg-H	73.6
5	Bz-**N-Me-Nle**-Lys-Arg-Arg-H	9.3
6	Bz-Nle-Lys-Arg-D-**Arg**-H	14.0
7	Bz-Nle-Lys-D-**Arg**-Arg-H	128.6
8	Bz-Nle-D-**Lys**-Arg-Arg-H	23.8
9	Bz-D-**Nle**-Lys-Arg-Arg-H	2.6
10	Bz-Nle-Lys-Arg-**Ala**-H	4.6
11	Bz-Nle-Lys-**Ala**-Arg-H	262.0
12	Bz-Nle-**Ala**-Arg-Arg-H	3.8
13	Bz-**Ala**-Lys-Arg-Arg-H	0.7
14	Bz-Nle-Lys-Arg-**Phe**-H	109.8
15	Bz-Nle-Lys-**Phe**-Arg-H	108.0
16	Bz-Nle-**Phe**-Arg-Arg-H	4.2
17	Bz-**Phe**-Lys-Arg-Arg-H	1.2
18	Bz-Nle-Lys-Arg-**Phg**-H	90.9
19	Bz-Nle-Lys-Arg-**(*p*-Ph)-Phe**-H	22.7
20	Bz-Nle-Lys-Arg-**(*p*-CN)-Phe**-H	62.0
21	Bz-Nle-Lys-Arg-**Tyr**-H	41.5
22	Bz-Nle-Lys-Arg-**His**-H	43.1
23	Bz-Nle-Lys-Arg-**Lys**-H	57.7
24	Bz-Nle-Lys-Arg-**Trp**-H	10.0
25	Bz-Nle-Lys-Arg-**(*p*-guanidinyl)-Phe**-H	11.8
26	Bz-Nle-Lys-Arg-**(*p*-Cl)-Phe**-H	202.8
27	Bz-Nle-Lys-**Lys(Z)**-Arg-H	99.5
28	Bz-Nle-Lys-**Lys**-Arg-H	1.9

[a] Bold residues highlight difference with respect to the reference tetrapeptide 1.

Table 2. Amino Acid Residues Forming the S1−S4 Pockets of the West Nile Virus (WNV) in the Bz-nKRR−WNV NS2B/NS3 Crystal Structure (2FP7)

pocket	NS2B	NS3
S1[a]	n/a	D129, Y130, P131, T132, S135, Y150, G151, Y161
S2[b]	G83,N84, F85	H51, E74, D75, N152
S3[c]	F85, Q86, L87	G153, V154, I155
S4	n/a	I155

[a] Residues predicted to be in the S1 pocket by Fairlie et al.:[17] V115, D129, Y150, S163, I165. Those predicted by Young et al.:[16] V115, D129, Y150, S160, S163. [b] Residues predicted to be in the S2 pocket by Fairlie et al.:[17] H51, D75, G151, N152. Those predicted by Young et al.:[16] N152. [c] Residues predicted to be in the S3 pocket by Fairlie et al.:[17] G153, I155.

water molecules, except O22, which probably plays an important role in the conformation adopted by the Bz-nKRR tetrapeptide aldehyde (see discussion below), were removed. The complexes were optimized using default parameters in Macromodel 9.0,[25] except relaxation was allowed to continue for at most 5000 cycles using the MMFFs[26,27] force field with the PRCG[28] optimization scheme. In all calculations, the tetrapeptide inhibitors and any amino acid residue having an atom within 8 Å of the inhibitor were allowed to relax while the remainder of the protein was kept frozen. When the procedure outlined above was benchmarked against the structure of the Bz-nKRR−WNV NS2B−NS3pro complex (2FP7), the original crystal structure was closely reproduced, thus suggesting that the procedure is suitable for the modeling study conducted here.

Discussion

Detailed analysis of the folds adopted in the Bz-nKRR−WNV NS2B/NS3 complex (2FP7) has been presented previously, with the inhibitor−protein interaction briefly discussed.[20] Further inspection of the X-ray crystal structure revealed important information about the nature of the binding pockets (Table 2). The S1 pocket is the most well-defined binding pocket in the crystal structure. With the carbonyl carbon of the P1 arginine bonded covalently to the catalytic S135 (C−O = 1.36 Å), the S1 pocket is predominantly hydrophobic due to the presence of two tyrosine side chains (NS3: Y150, Y161) forming the bottom and the edge of the pocket, with opportunities for substrate/inhibitor−enzyme hydrogen bonding via the protein backbone atoms of NS3 (Y130, P131, S135 and G151) and the side chain of NS3 (D129 and Y151) (Figure 1). Even though the S2 pocket is also well-defined, it is much narrower when compared to S1. Lined with hydrogen-bond donors and acceptors, the S2 pocket is defined by the NS2B (G83, N84, and F85) and NS3 (H51, E74, D75, and N152) residues (Figure 1). In contrast to S1 and S2, the S3 pocket is relatively wide and shallow, with one side of the pocket lined with hydrophobic side chains (NS3: V154, I155), and the other side comprised of hydrogen-bond donor/acceptor atoms from the backbone of NS2B (F85, Q86, L87) and NS3 (G153) (Figure 1). The vague shape and contrasting binding environment on the sides of the pocket lead to a rather undiscriminating S3 pocket. The binding region of S4, made up entirely of the solvent-exposed hydrophobic residue I155, might be better described as a "patch", since the shape of the pocket is ill-defined. Overall, the crystal structure suggests that the S1 and S2 pockets would be the dominant peptide recognition sites, which is in agreement with the homology model of the WNV NS2B/NS3 recently published by Young et al.[16] However, the importance of the NS2B cofactor contribution to the S2 and S3 pockets was not predicted in their work. Interestingly, while their homology model suggested that V115 forms part of the S1 pocket, in the crystal structure 2FP7, the residue is located some 18 Å from the pocket. Given this observation, it is unlikely that the inactivation of the WNV protease by V115 mutation is a consequence of *direct* perturbation of the S1 pocket as originally proposed.[16] In the crystal structure, this valine is located in the B2b region of the NS3 domain, with the backbone of the residue interacting with E73 and V75 of the NS2B β3 strand. Even though the backbone−backbone interaction would be retained with V115A and V115F mutation, it is plausible that changes in the side chain may perturb the region sufficiently to disturb the essential interface between NS2B and NS3. With V115 located within 4 Å of NS2B R74, V → F mutation creates an opportunity for cation−π interaction between the side chains. This would lead to changes in the protein, which provides an alternative rationale for the inactivation of the protease upon V115F mutation.

In the crystal structure 2FP7, the P1 arginine in Bz-nKRR aldehyde is stabilized by interaction between the carbonyl oxygen and H51, hydrogen bonding with side chain interactions with D129 and Y130, and a backbone interaction with G151 (Figure 1). On the basis of a homology model, Young et al.[16] predicted that P1−R would interact with S160 and S163 in the S1 pocket. However, in the crystal structure, these two serine residues are not found in the S1 pocket. Thus, the observed decrease in substrate affinity upon S160A and S163A mutation may not be due to the perturbation of hydrogen bonds to the P1−R as originally suggested. Here we noted that the side chain oxygen of S163 and S160 are hydrogen bonded to G151 and N158, respectively. As G151 forms part of the S1 pocket, adjacent to the important residue Y150, disruption of hydrogen bonding upon S → A mutation will change the shape/size of the S1 pocket directly. Previously, it has been noted that one of the key differences between the *liganded* WNV and the *unliganded* Dengue virus crystal structure lies in the NS2B folds beneath the β-loop E2b−F2.[20] With the N158 residue located at the loop between E2b and F2 β-sheets, S160A mutation is very likely to change the shape/size of the S1 pocket indirectly. In other words, even though S160 and S163 are not in the S1 pocket and hence not interacting with P1−R as proposed,[16] the

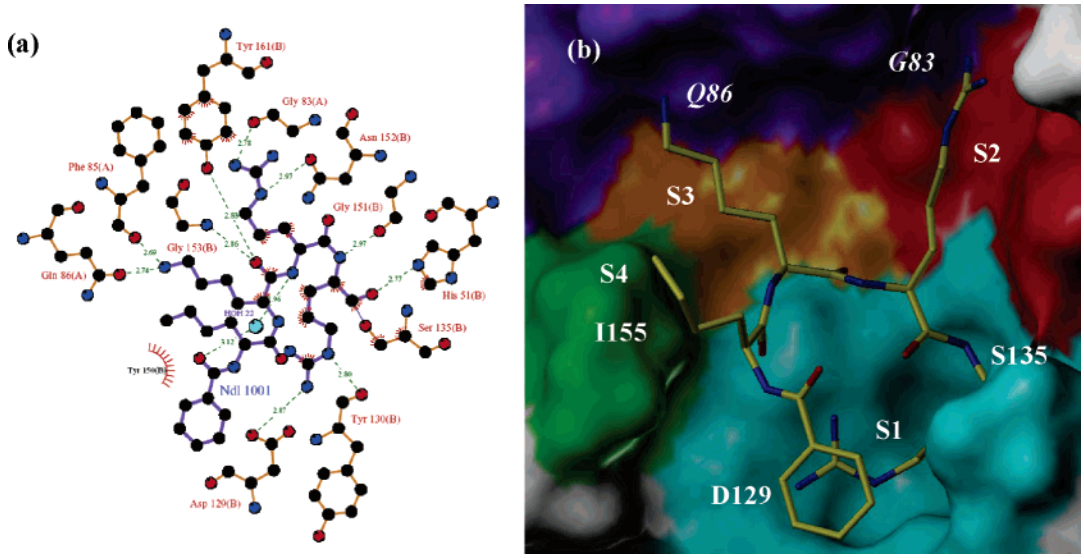

Figure 1. (a) A two-dimensional plot of the key interactions between Bz-nKRR and WNV protease in 2FP7, with the contributions from NS2B and NS3 domain labeled as A and B, respectively. (b) Connelly surface of the WNV active site with the Bz-nKRR inhibitor shown in stick format with hydrogen removed for clarity; residues contributing to the S1, S2, S3, and S4 pockets colored as cyan, red, orange, and green, respectively. The NS2B contributions to the active site (S2 and S3) are colored purple. Note the phenyl ring of the P4-Bz cap lies almost parallel to the positively charged P1-R side chain.

shape/size of the S1 pocket would be perturbed upon S → A mutation, thus providing an alternative rationale to the observed decreased in substrate affinity. The P2−S2 interaction comprises two hydrogen bonds, with the arginine interacting with NS3 N152 and NS2B G83 (Figure 1), while the P3−K makes four hydrogen bonds: two from the lysine side chain to NS2B F85 and Q86 and another two between the backbone carbonyl oxygen of lysine with the backbone amide of G153 and the side chain of Y161. No hydrogen bonding is observed between the P4-norleucine and the S4 binding region.

Interestingly, the Bz-nKRR peptide inhibitor (**1**) adopts a rather unusual "cyclic" conformation, in which the phenyl ring of the P4−Bz cap lies almost parallel to the positively charged P1−R side chain so that the guanidine is sandwiched between the phenyl rings of the Bz cap and Y150. This conformation is stabilized by three factors: cation−π interaction between Bz cap and the guanidine side chain (~3.49 Å from the centroid of the phenyl-π ring to the guanidine carbon); the presence of a bridging water between P1−R (carboxyl O···O22 = 3.39 Å), P2−R (amide N···O22 = 2.96 Å), and P4−Bz cap (carboxyl O···O22 = 3.12 Å); and an intramolecular hydrogen bond between the P1−R side chain and the P2 carbonyl oxygen (N···O = 3.10 Å). A proline mutation was performed in the P3 position to assess the characteristics of the inhibitor conformation. Despite losing the side chain hydrogen bonds (to F85 and Q86), as the peptide Bz-nPRR (**2**) retains the backbone conformation of Bz-nKRR (**1**) (Figure 2), the peptide remains moderately potent. This thus suggests an opportunity for adding rigidity to the design of potent inhibitors.

One of the characteristics of the serine protease molecular recognition mechanism is the backbone−backbone hydrogen bonds made with the substrate. Thus, a methyl scan of the backbone amide nitrogen was performed to probe this aspect. Methylation of the amide nitrogen of P4 (**5**) had little effect on the IC_{50}, which is expected, as the NH was not interacting with the enzyme in the crystal structure. The only notable geometrical change is a slight shift of the P4−Bz cap, which probably leads to a diminished cation−π interaction with the P1−R and may account for the 2-fold decrease in IC_{50}. Methylating the P3 amide nitrogen (**4**) led to an 18-fold loss of potency of the

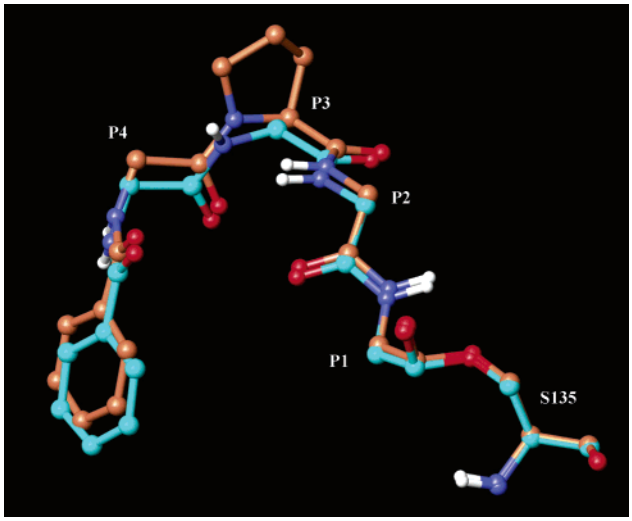

Figure 2. Model of Bz-nPRR (**2**, coral) in WNV protease and Bz-nKRR (**1** from 2FP7, cyan), with nonpolar hydrogens and side chains removed for clarity. Note the backbone conformation of the peptide found in 2FP7 is well-retained when the P3 lysine is replaced by proline.

tetrapeptide. Our model suggests that this peptide binds to WNV protease in an identical fashion as Bz-nKRR, so this drastic loss of potency may be related to some dynamic process of molecular recognition. Last, methylating the P2 amide nitrogen (**3**) decreased the peptide inhibitor activity by nearly 14-fold. This may be due to the displacement of the bridging water (O22) in the crystal structure (arising from the loss of the amide NH···O22 interaction), which was postulated to be partly responsible for stabilizing the tetrapeptide in the "cyclic" conformation.

A stereoscan of the reference tetrapeptide Bz-nKRR was also performed to gain a better understanding of the stereoselectivity of the binding pockets. Given the nature of the S4 binding region, substituting the P4-norleucine$_L$ with norleucine$_D$ (**9**) as expected did little to change the activity of the peptide. Surprisingly, the P1−R_L to R_D substitution (**6**) also had little effect. Overlapping our model for **6** with the crystal structure, we found that the side chain guanidinyl group at P1 was in a similar position for both peptides (Figure 3a), thus not disrupting

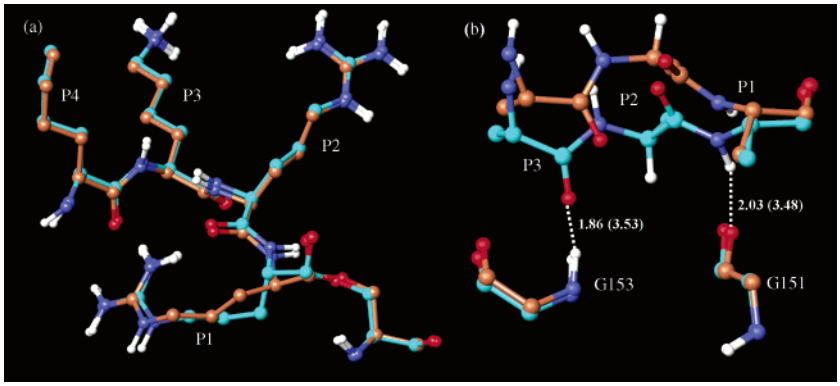

Figure 3. (a) Model of Bz-nKRR$_D$ (**6**, coral) in WNV protease and Bz-nKRR$_L$ (**1** from 2FP7, cyan). Note the very good overlap between the two tetrapeptides, except for the minor deviation for C_β and C_γ. The P4-Bz cap, nonpolar hydrogens of the peptides, and all protein atoms are not shown for clarity. (b) Model of Bz-nKR$_D$R (**7**, coral) in WNV protease and Bz-nKR$_L$R (**1** from 2FP7, cyan), showing only the backbone atoms and polar hydrogens of P1−P3 along with G151 and G153 of WNV NS3. The distance, in Å, for the pairs of backbone hydrogen bonds found in 2FP7, 1.86 Å (2.76 Å O−N) and 2.03 Å (2.97 Å N−O), with the corresponding distance in the model for Bz-nKR$_D$R (**7**)−WNV, 3.53 Å (4.43 Å O−N) and 3.48 Å (3.94 Å O−N), included in parentheses.

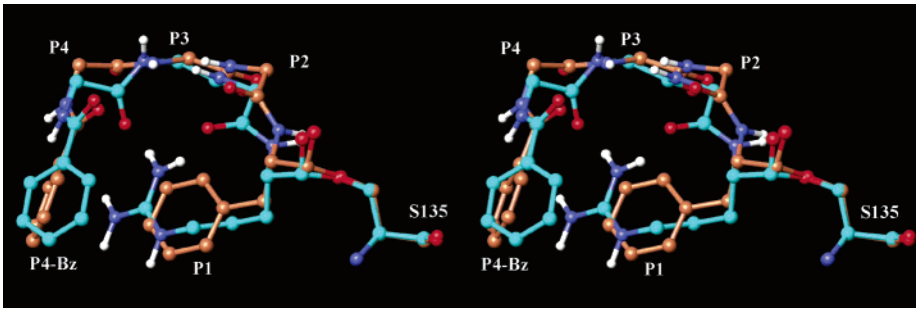

Figure 4. Stereoview for the model of Bz-nKRF (**14**, coral) in WNV protease and Bz-nKRR$_L$ (**1** from 2FP7, cyan): note that in Bz-nKRF, the P4-Bz cap and the P1-F are not stacked.

the important interactions found in the Bz-nKRR (**1**). Our model suggests that regardless of the stereochemistry for the P1 arginine, the interaction with D129 in the S1 pocket is retained. Changing the stereocenter of P3 (lysine$_L$ to lysine$_D$, **8**) led to a 7-fold decrease in potency. The root cause of the potency loss is not clear from our current model. Converting P2-R$_L$ to R$_D$ (**7**) led to a loss in potency of more than 30-fold. Our model suggests that such substitution may lead to the disruption of backbone hydrogen bonds between the tetrapeptide and the protein, with the distance between P1-R amide nitrogen with G151 and P3 carbonyl oxygen with G153 increasing by over 1.5 Å when compared to that of 2FP7 (Figure 3b).

To gain insight into the WNV protease substrate binding characteristics, an alanine scan of the reference compound Bz-nKRR (**1**) was performed. The IC_{50s} suggests that potency of the tetrapeptide is unaffected with substitution at the P3 (**12**) position. This is to be expected for the P3 position, as the backbone hydrogen-bonding interaction is retained when the K is substituted with A. Replacing the P1- or P2-Arg by alanine (**10**, **11**) led to a dramatic loss of potency. This is also expected given the importance of the S1 and S2 pockets in substrate recognition, and the loss of side chain interactions (D129, T130 for **10** and N152 for **11**). Interestingly, alanine substitution at the P4 position (Bz-AKRR, **13**) yielded a slightly more potent peptide. Given the general lack of interaction between S4 and P4, this change in potency is probably driven by solvent entropy when the longer hydrophobic norleucine is being replaced by the shorter alanine side chain.

We further probed the importance of π residues in the binding pockets with a phenylalanine scan of the reference compound. Results from the substitution of P2 (**15**), P3 (**16**), and P4 (**17**) are consistent with our discussions presented above. Substitution at the P1 position with a phenylalanine (Bz-nRKF, **14**) resulted in a much less potent inhibitor when compared to **1**. This was rather surprising given that the WNV S1 pocket is lined with aromatic residues. Our model suggests that the P1 phenyl ring of Bz-nRKF may not be extending far enough into the S1 pocket to take advantage of the π stacking environment presented by the Bz cap (Figure 4). In addition, the hydrogen bonding between the guanidinyl group and D129 was lost upon phenylalanine substitution, so the loss of potency of Bz-nKRF (**14**) may be due to a combination of both factors.

In an effort to exploit the aromatic character and hydrogen bonding available in the S1 pocket, a series of peptides was synthesized and evaluated (**18**−**26**). Replacing the P1 phenylalanine with phenylglycine (**18**) led to a similarly inactive tetrapeptide. On the other hand, adding a phenyl ring (**19**) or cyano group (**20**) at the para-position of phenylalanine increased the activity over Bz-nKRF (**14**). These results are consistent with our hypothesis that the phenyl group in Bz-nKRF (and hence also in **18**) is too far from Y161 to take advantage of the π stacking environment present in the S1 pocket, which can be improved by extending the π system in the case of **19** and **20**. Tyrosine (**21**), which has a π ring at the same location as phenylalanine in **14**, had a 2-fold improvement in activity, presumably due to hydrogen bonding via the phenol OH to the side chain of D129. Similarly, the NH_3^+ group of lysine in Bz-nKRK (**23**) also forms a hydrogen bond with the side chain of D129, thus yielding a comparable IC_{50} to **21**. However, Bz-nKRY [likewise Bz-nKRH (**22**) and Bz-nRKK (**23**)] was still 10-fold less active than Bz-nKRR due to either the loss of the $\pi-\pi$ interactions or hydrogen bonds. Both criteria are met in Bz-nKRW (**24**): with the combined effect of $\pi-\pi$ interaction (the indole ring with Y161 and Bz capping group) and hydrogen

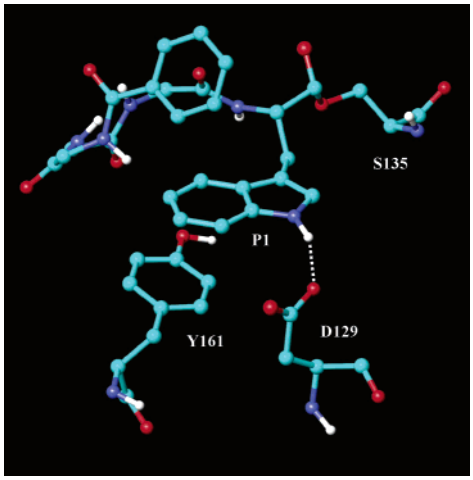

Figure 5. Model of Bz-nKRW (**24**) in WNV protease, with P2−P4 side chains and nonpolar hydrogens removed for clarity. Note the $\pi-\pi$ stacking between the Y161 of the NS3 Pro, the imidizole group of the P1-W, and the P4-Bz cap. The hydrogen bonding (highlighted with the dotted line) distance between P1-W and D129 is 1.76 Å.

bonding (with D129) (Figure 5), substitution of the P1 with tryptophan afforded a substantial (nearly 10-fold) increase in activity over that of Bz-nKRF (**14**). Interestingly, a guanidinyl group at the para-position of the phenyl ring (**25**) afforded an activity comparable to the P1 substitution of tryptophan (**24**). Our model suggests that even though the phenyl group of **25** contributes very little to the binding to the enzyme, the guanidinyl group is interacting with the side chain of D129, via a pair of hydrogen bonds, thus explaining the potency of **25**. Last, the chloro substitution at the para-position of the phenylalanine (**26**) decreased the activity 2-fold over that of Bz-nKRF, which could be explained by the unfavorable electrostatic repulsion between the bulky and electronegative chlorine atom with the carboxylate group on D129. In fact, the change in the protein structure caused by the chlorine atom is quite significant, with the carboxylate group of D129 shifted by ~0.8 Å when the model of **26** is compared to that of **14**.

Finally, we turned our attention to the SAR of the P2−S2 interaction probed by our tetrapeptide inhibitors. When the P2 arginine was replaced by Lys(Z) (**27**), a 25-fold reduction in potency was observed. This is consistent with the fact that the S2 pocket is relatively shallow and thus would have difficulty accommodating the rather large Lys(Z) moiety. On the other hand, the activity of Bz-nKKR (**28**) was comparable to that of **1**, as the key hydrogen bonding with D75 is retained, and the lysine terminal NH_3^+ makes a comparable interaction with the WNV NS2B domain (G83) to the P2-R in 2FP7. On the basis of homology models and apparently supported by a mutagenesis study (N152A), it has been suggested that hexapeptide substrates/inhibitors with lysine at the P2 position would be hydrogen bonded to the N152 residue in the S2 pocket.[16,17] Given the flexibility of lysine, it is plausible that it could interact with N152. However, it is of interest to note that the interaction between the P2 lysine and G83 would not have been predicted without the recent crystal structure, given that the location of the NS2B domain would be predicted differently in a homology model. Furthermore, as the crystal structure suggests that the side chain of N152 is hydrogen bonded to G83, the result of the mutagenesis study of N152A could be due to a change in size/shape of the S2 pocket, rather than to the direct interaction between N152 and P2-K.

Conclusion

For the first time, a modeling study has been conducted using the crystal structure of a catalytically active form of WNV NS3 protease for a series of substrate-based tetrapeptide aldehydes. We found that the crystal structure is relevant in explaining the observed SAR for this series of tetrapeptides, confirming the importance of the NS2B domain in substrate-based inhibitor binding of WNV. In agreement with previous studies based on homology models of the WNV protease, we found that the S1 and S2 pockets are the key peptide recognition sites. However, some residues (e.g., V115, S160, S163) that were suggested to form part of these pockets in the homology models were found to be distant from these two sites in the crystal structure. To explain the loss of substrate activity arising from mutating these residues, we have proposed alternative rationales in this study. In general, a residue with an opportunity for hydrogen bonding and π-stacking is favored in the S1 pocket, while a positively charged residue is preferred in the S2 pocket. With the success of the present study, it also suggests that the WNV NS2B/NS3 protease crystal structure employed here would provide useful guidance in the drug discovery process for related *Flavivirus* proteases, given the large degree of homology within the family.[15]

References

(1) Smithburn, K. C.; Hughs, T. P.; Burke, A. W.; Paul, J. H. A neurotropic virus isolated the blood of a native of Uganda. *Am. J. Trop. Med. Hyg.* **1940**, 471−492.
(2) Monath, T. P. Flaviviruses. *Virology*; Raven Press: New York, 1990; pp 763−814.
(3) Jia, X. Y.; Briese, T.; Jordan, I.; Rambaut, A.; Chi, H. C.; Mackenzie, J. S.; Hall, R. A.; Scherret, J.; Lipkin, W. I. Genetic analysis of West Nile New York 1999 encephalitis virus. *Lancet* **1999**, *354*, 1971−1972.
(4) Taylor, R. M.; Work, T. H.; Hurlbut, H. S.; Rizk, F. A study of the ecology of West Nile Virus in Egypt. *Am. J. Trop. Med. Hyg.* **1956**, *5*, 579−620.
(5) Philip, C. B.; Smadel, J. E. Transmission of West Nile virus by infected *Aedes albopictus*. *Proc. Soc. Exp. Biol. Med.e* **1943**, *53*, 49−50.
(6) Petersen, L. R.; Marfin, A. A. West Nile virus: A primer for the clinician. *Ann. Intern. Med.* **2002**, *137*, 173−179.
(7) Lanciotti, R. S.; Roehrig, J. T.; Deubel, V.; Smith, J.; Parker, M.; Steele, K.; Crise, B.; Volpe, K. E.; Crabtree, M. B.; Scherret, J. H.; Hall, R. A.; MacKenzie, J. S.; Cropp, C. B.; Panigrahy, B.; Ostlund, E.; Schmitt, B.; Malkinson, M.; Banet, C.; Weissman, J.; Komar, N.; Savage, H. M.; Stone, W.; McNamara, T.; Gubler, D. J. Origin of the West Nile virus responsible for an outbreak of encephalitis in the northeastern United States. *Science* **1999**, *286*, 2333−2337.
(8) Perni, R. B.; Almquist, S. J.; Byrn, R. A.; Chandorkar, G.; Chaturvedi, P. R.; Courtney, L. F.; Decker, C. J.; Dinehart, K.; Gates, C. A.; Harbeson, S. L.; Heiser, A.; Kalkeri, G.; Kolaczkowski, E.; Lin, K.; Luong, Y. P.; Rao, B. G.; Taylor, W. P.; Thomson, J. A.; Tung, R. D.; Wei, Y. Y.; Kwong, A. D.; Lin, C. Preclinical profile of VX-950, a potent, selective, and orally bioavailable inhibitor of hepatitis C virus NS3−4A serine protease. *Antimicrob. Agents Chemother.* **2006**, *50*, 899−909.
(9) Tomei, L.; Failla, C.; Santolini, E.; Defrancesco, R.; Lamonica, N. Ns3 is a serine-protease required for processing of hepatitis-C virus polyprotein. *J. Virol.* **1993**, *67*, 4017−4026.
(10) Eckart, M. R.; Selby, M.; Masiarz, F.; Lee, C.; Berger, K.; Crawford, K.; Kuo, C.; Kuo, G.; Houghton, M.; Choo, Q. L. The hepatitis-C virus encodes a serine protease involved in processing of the putative nonstructural proteins from the viral polyprotein precursor. *Biochem. Biophys. Res. Commun.* **1993**, *192*, 399−406.
(11) Malcolm, B. A.; Liu, R.; Lahser, F.; Agrawal, S.; Belanger, B.; Butkiewicz, N.; Chase, R.; Gheyas, F.; Hart, A.; Hesk, D.; Ingravallo, P.; Jiang, C.; Kong, R.; Lu, J.; Pichardo, J.; Prongay, A.; Skelton, A.; Tong, X.; Venkatraman, S.; Xia, E.; Girijavallabhan, V.; Njoroge, F. G. SCH 503034, a mechanism-based inhibitor of hepatitis C virus NS3 protease, suppresses polyprotein maturation and enhances the antiviral activity of alpha interferon in replicon cells. *Antimicrob. Agents Chemother.* **2006**, *50*, 1013−1020.

(12) Chambers, T. J.; Grakoui, A.; Rice, C. M. Processing of the yellow-fever virus nonstructural polyprotein−A catalytically active Ns3-proteinase domain and Ns2b are required for cleavages at dibasic sites. *J. Virol.* **1991**, *65*, 6042−6050.
(13) Falgout, B.; Pethel, M.; Zhang, Y. M.; Lai, C. J. Both nonstructural proteins Ns2b and Ns3 are required for the proteolytic processing of dengue virus nonstructural proteins. *J. Virol.* **1991**, *65*, 2467−2475.
(14) Shiryaev, S. A.; Ratnikov, B. I.; Chekanov, A. V.; Sikora, S.; Rozanov, D. V.; Godzik, A.; Wang, J.; Smith, J. W.; Huang, Z. W.; Lindberg, I.; Samuel, M. A.; Diamond, M. S.; Strongin, A. Y. Cleavage targets and the d-arginine-based inhibitors of the West Nile virus NS3 processing proteinase. *Biochem. J.* **2006**, *393*, 503−511.
(15) Ganesh, V. K.; Muller, N.; Judge, K.; Luan, C. H.; Padmanabhan, R.; Murthy, K. H. M. Identification and characterization of nonsubstrate based inhibitors of the essential Dengue and West Nile virus proteases. *Bioorg. Med. Chem.* **2005**, *13*, 257−264.
(16) Chappell, K. J.; Nall, T. A.; Stoermer, M. J.; Fang, N. X.; Tyndall, J. D. A.; Fairlie, D. P.; Young, P. R. Site-directed mutagenesis and kinetic studies of the West Nile virus NS3 protease identify key enzyme−substrate interactions. *J. Biol. Chem.* **2005**, *280*, 2896−2903.
(17) Nall, T. A.; Chappell, K. J.; Stoermer, M. J.; Fang, N. X.; Tyndall, J. D. A.; Young, P. R.; Fairlie, D. P. Enzymatic characterization and homology model of a catalytically active recombinant West Nile virus NS3 protease. *J. Biol. Chem.* **2004**, *279*, 48535−48542.
(18) Love, R. A.; Parge, H. E.; Wickersham, J. A.; Hostomsky, Z.; Habuka, N.; Moomaw, E. W.; Adachi, T.; Margosiak, S.; Dagostino, E.; Hostomska, Z. The conformation of hepatitis C virus NS3 proteinase with and without NS4A: A structural basis for the activation of the enzyme by its cofactor. *Clin. Diagn. Virol.* **1998**, *10*, 151−156.
(19) Leung, D.; Schroder, K.; White, H.; Fang, N. X.; Stoermer, M. J.; Abbenante, G.; Martin, J. L.; Young, P. R.; Fairlie, D. P. Activity of recombinant dengue 2 virus NS3 protease in the presence of a truncated NS2B cofactor, small peptide substrates, and inhibitors. *J. Biol. Chem.* **2001**, *276*, 45762−45771.
(20) Erbel, P.; Schiering, N.; D'Arcy, A.; Renatus, M.; Kroemer, M.; Lim, S. P.; Yin, Z.; Keller, T. H.; Vasudevan, S. G.; Hommel, U. Structural basis for the activation of flaviviral NS3 proteases from dengue and West Nile virus. *Nat. Struct. Mol. Biol.* **2006**, *13*, 372−373.
(21) Yin, Z.; Patel, S. J.; Wang, W. L.; Chan, W. L.; Rao, K. R. R.; Wang, G.; Ngew, X.; Patel, V.; Beer, D.; Knox, J. E.; Ma, N. L.; Ehrhardt, C.; Lim, S. P.; Vasudevan, S. G.; Keller, T. H. Peptide inhibitors of dengue virus NS3 protease. Part 2: SAR study of tetrapeptide aldehyde inhibitors. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 40−43.
(22) Yin, Z.; Patel, S. J.; Wang, W. L.; Wang, G.; Chan, W. L.; Rao, K. R. R.; Alam, J.; Jeyaraj, D. A.; Ngew, X.; Patel, V.; Beer, D.; Lim, S. P.; Vasudevan, S. G.; Keller, T. H. Peptide inhibitors of dengue virus NS3 protease. Part 1: Warhead. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 36−39.
(23) *SYBYL 6.9*; Tripos Inc., 1699 South Hanley Road., St. Louis, MO 63144.
(24) *Maestro 7.0*; Schrodinger, LLC, New York, NY, 1999−2005.
(25) *MacroModel*, version 9.0; Schrodinger, LLC, New York, NY, 2005.
(26) Halgren, T. A. MMFF VII. Characterization of MMFF94, MMFF94s, and other widely available force fields for conformational energies and for intermolecular-interaction energies and geometries. *J. Comput. Chem.* **1999**, *20*, 730−748.
(27) Halgren, T. A. MMFF VI. MMFF94s option for energy minimization studies. *J. Comput. Chem.* **1999**, *20*, 720−729.
(28) Polak, E.; Ribiere, G. Note on convergence of conjugate direction methods. *Rev. Fr. D Inform. Rech. Oper.* **1969**, *3*, 35.

JM0607606