Biochemical and Biophysical Research Communications 434 (2013) 767–772

Contents lists available at SciVerse ScienceDirect

Biochemical and Biophysical Research Communications

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

Design and evaluation of substrate-based octapeptide and non substrate-based tetrapeptide inhibitors of dengue virus NS2B–NS3 proteases

Peteris Prusis [a], Muhammad Junaid [a], Ramona Petrovska [a], Sviatlana Yahorava [a], Aleh Yahorau [a], Gerd Katzenmeier [b], Maris Lapins [a], Jarl E.S. Wikberg [a,*]

[a] Department of Pharmaceutical Biosciences, Division of Pharmacology, Uppsala University, 75124 Uppsala, Sweden
[b] Laboratory of Molecular and Cellular Microbiology, Institute of Molecular Biosciences, Mahidol University, 73170 NakornPathom, Thailand

ARTICLE INFO

Article history:
Received 26 March 2013
Available online 12 April 2013

Keywords:
Dengue virus
NS2B–NS3 protease
D-optimal design
Substrate-based
Non-substrate-based
Tetrapeptide inhibitor

ABSTRACT

A series of 45 peptide inhibitors was designed, synthesized, and evaluated against the NS2B–NS3 proteases of the four subtypes of dengue virus, DEN-1–4. The design was based on proteochemometric models for Michaelis (K_m) and cleavage rate constants (k_{cat}) of protease substrates. This led first to octapeptides showing submicromolar or low micromolar inhibitory activities on the four proteases. Stepwise removal of cationic substrate non-prime side residues and variations in the prime side sequence resulted finally in an uncharged tetrapeptide, WYCW-NH_2, with inhibitory K_i values of 4.2, 4.8, 24.4, and 11.2 μM for the DEN-1–4 proteases, respectively. Analysis of the inhibition data by proteochemometric modeling suggested the possibility for different binding poses of the shortened peptides compared to the octapeptides, which was supported by results of docking of WYCW-NH_2 into the X-ray structure of DEN-3 protease.

1. Introduction

Mosquito-borne dengue virus belongs to the genus *Flavivirus* that comprises over 70 viruses many of which are human pathogens such as West Nile virus, yellow fever virus, and Japanese encephalitis virus. Infection by dengue virus causes symptoms from mild fever to life-threatening dengue hemorrhagic fever and dengue shock syndrome. The virus is endemic to most tropical and subtropical regions, with more than 2.5 billion people at risk for epidemic transmission, and an estimated 50 million infections each year. There are four highly homologous but antigenically distinct serotypes of the virus (DEN-1–4), and immunity against one serotype does not protect against infection by another serotype; instead it may enhance disease severity upon infection with another serotype [1].

Currently there are no antiviral drugs available against flaviviruses. An attractive target for development of therapeutic inhibitors is the two-component viral protease NS2B–NS3. The flaviviral NS3 protein is a multifunctional enzyme that possesses protease, helicase and RNA triphosphatase activities. The N-terminal domain of about 180 amino acids of NS3 is a trypsin-like serine protease, which adopt a chemotrypsin-like fold with two β-barrels, each formed by six β-strands, with the catalytic triad His51-Asp75-Ser135 located at the cleft between the β-barrels. For full enzymatic activity the protease requires the hydrophilic part of the integral membrane protein NS2B (residues 49–95). Several X-ray crystal structures of dengue and West Nile virus proteases have been determined. In all of them the NS2B cofactor is located at a distance from the substrate binding site in ligand-free protease and is lining the substrate binding site in the inhibitor-bound protease. In the latter conformation, NS2B provides a β-strand to the N-terminal β-barrel and stabilizes this domain by covering hydrophobic residues of NS3 [2–4].

Considerable efforts have been recently made to find inhibitors to proteases of dengue virus, West Nile virus, and other flaviviruses. Peptide inhibitors for dengue NS2B–NS3 showing low micromolar activities were recently reported [5,6]. Such inhibitors were based on non-prime side sequences of substrate cleavage sites; all which contain basic amino acids at the P1 and P2 positions and often also at the P3 and P4 positions [7]. High charge is associated with poor membrane permeability, however, making non-prime side sequences unattractive as antivirals. This is because viral replication occurs intracellularly, and without cell-penetration the compound will be ineffective.

Small molecule inhibitors of dengue NS2B–NS3 reaching low micromolar activities were also reported, which included anthracene [8] and phthalazine-based compounds [9], and flavonoids

Abbreviations: DEN-1–4, dengue virus serotypes 1–4; PCM, proteochemometric; PLS, partial least squares regression; Abz, *o*-aminobenzoic acid; nY, 3-nitrotyrosine.

* Corresponding author. Fax: +46 18 55 9718.
E-mail address: Jarl.Wikberg@farmbio.uu.se (J.E.S. Wikberg).

0006-291X
http://dx.doi.org/10.1016/j.bbrc.2013.03.139

[10]. Non-prime side peptide sequences, and compounds based on aromatic scaffolds were also found, which inhibited NS2B–NS3 proteases of West Nile and yellow fever viruses [11–14].

In a previous study we investigated the prime side specificity of dengue virus NS2B–NS3 protease substrates [15]. Using statistical experimental design we selected a set of 48 internally quenched peptides with the general structure Abz-RRRR|XXXX-nY-NH_2 (where Abz is *o*-aminobenzoic acid and nY is 3-nitrotyrosine). Thus, the non-prime side sequence of all these peptides corresponded to the sequence of the cleavage site of DEN-2 capsid protein, while for the P′1–P′4 positions 6–10 amino acids were varied for each position under the constraint of D-optimal design [16]. The peptides were synthesized and assayed with DEN-1–4 NS2B–NS3 proteases for Michaelis (K_m) and cleavage rate constants (k_{cat}). The results were analyzed by proteochemometric (PCM) modeling, which showed that different physico-chemical properties of amino acids contributed independently to the two activities. It was e.g. found that the presence of hydrophobic or rigid amino acids at the P′1 position is disadvantageous for substrate cleavage, while on the other hand they exert less influence on substrate binding. Interpretation of the PCM model also suggested that a cysteine in the P′3 position gives better binding (i.e. lower K_m) than any of the amino acids present at P′3 in dengue virus' native cleavage sites. Moreover, most of the peptides with C in the P′3 or P′4 positions showed substrate inhibition, and furthermore a peptide with the prime side sequence GWCF was non-cleavable by all the DEN-1–4 proteases. These findings indicated the possibility for design of new peptides acting as inhibitors of dengue proteases. In this study we exploited this finding, which resulted in an uncharged peptide with micromolar inhibitory activity on the NS2B–NS3 proteases from four different dengue serotypes.

2. Materials and methods

2.1. Expression and purification of dengue virus proteases

We used our previously described pTrcHis plasmids containing dengue virus protease sequences [17]. Expression of NS2B(H)–NS3pro from these plasmids was done in *Escherichia coli*. Dengue virus type 2–4 proteases were purified from inclusion bodies and refolded, as described previously [17]. Dengue virus type 1 protease was purified from supernatants, as described previously [18].

2.2. Synthesis of peptides

Forty-five peptides were prepared by solid-phase synthesis using an automated multiple peptide synthesizer (MultiPep; Intavis Bioanalytical Instruments AG, Koeln, Germany), using its automated standard protocol optimized for Fmoc chemistry as described previously [15]. All chemicals were reagent grade from Fluka (Sigma–Aldrich, St. Louis, MO, USA), Applied Biosystems (Foster City, CA, USA), Bachem (Bubendorf, Switzerland) and Novabiochem (Calbiochem-Novabiochem AG, Laufelfingen, Switzerland). Peptides were characterized by HPLC and their structures were confirmed by MS as described [15]. The purity of raw peptides was above 80%. After small-scale preparative HPLC the purity of all peptides was above 95% and they were used for inhibitor assays after freeze drying as described [15].

2.3. Inhibitor assay

The fluorogenic assay was carried out with the fluorogenic substrate Abz-RRRRSAG-nY-NH_2 on a FLUOstar OPTIMA 96-well plate reader (BMG Labtech GmbH), monitoring emission at 420 nm upon excitation at 320 nm. The final reaction volume was 100 μL and contained 50 mM Tris–HCl, pH 9.0 and 20% glycerol. Typically, test compounds were pre-incubated with 150 nM protease at 37 °C for 20 min. Substrate cleavage was then initiated by the addition of substrate at 10 μM final concentration. Reaction progress was monitored continuously during 60 min. Each test compound was diluted 1:2 from a 100 μM stock solution, with totally 11 dilution points for each test compound. For compounds that showed low activity the assay was repeated starting from 500 μM concentration. For controls only solvent was used. Inhibition constant K_i values were determined from dose–response curves using GraFit version 5 software.

2.4. Proteochemometric modeling of activity data

The inhibition K_i data for the 45 peptides for the four proteases was concomitantly analyzed by proteochemometric modeling (PCM), which aims at finding a quantitative relationship between the interaction data and descriptors of the interacting moieties [19].

For the PCM, the peptides were characterized by 14 descriptors. Five of these were binary and represented: (1) presence or absence of Abz and nY groups, (2) acetylation or non-acetylation of N-terminus, and (3–5) presence of at least one arginine, two arginines, or four arginines in the peptide. The remainder of the descriptors were obtained by characterizing each prime side residue with three quantitative physico-chemical *z*-scale descriptors [20], which gave totally nine descriptors for the varied P′1, P′2, and P′4 residues. (The invariant cysteine residue was viewed as to be located at the P′3 position for all peptides.) The three *z*-scales can be interpreted as representing hydrophilicity, steric properties, and polarity of amino acids. The four proteases were represented by four binary descriptors, where each protease was assigned the value 1 for one of the descriptors and 0 for the three others. All descriptors were mean centered and scaled to unit variance. The response variable (pK_i, i.e. negative logarithm of K_i value for DEN-1–4 protease inhibition) was also mean centered. For the case that a peptide did not show inhibitory activity, we assigned it $pK_i = 3$ for the sake of the PCM modeling (i.e. one millimolar activity).

The relationships between the descriptors and the response were established by partial least squares regression (PLS) using the OPLS algorithm as implemented in Simca-P+11 software (Umetrics AB). The optimal complexity of the PLS model was estimated by fivefold cross-validation, leaving out data for 1/5 of the peptides at a time. Thus, all data for any peptide (i.e. inhibition of DEN-1–4 proteases) were assigned to the same cross-validation group to avoid risk for model overfit for the case that a peptide shows similar activities against DEN-1–4. Overestimations of predictive performance and overfits may also occur if compounds used in model evaluation are very similar to the compounds in the model. To eliminate such risk, pairs of peptides that differed only by acetylation of the N-terminus were also included in the same cross validation group (e.g. WYCW-NH_2 was placed in the same group as Ac-WYCW-NH_2, etc.).

2.5. Molecular docking

The most active tetrapeptide found herein, WYCW-NH_2, was docked into the dengue virus protease using the recently solved crystal structure of DEN-3 NS2B–NS3 protease in complex with the aldehyde inhibitor Bz-nKRR-H (PDB code 3U1I) [4]. Prior to the docking the embedded inhibitor, water molecules and ions were removed from the PDB file, and polar hydrogen atoms were added using AutoDock Tools software [21]. The docking was performed by AutoDock Vina software [22]. The search space covered the whole NS2B–NS3 complex.

3. Results and discussion

3.1. Peptide library

In the first part of this study we designed and synthesized a set of substrate-like octapeptides with four arginines at the P4–P1 positions, and tested their inhibitory activity (K_i) on DEN-1–4 proteases. Subsequently, we studied changes on inhibitory activities caused by shortenings and other modifications of the peptides. Finally, we performed PCM analysis and 3D docking of the most active short peptide found into the DEN-3 NS2–NS3 protease complex.

The sequences and inhibitory activities of all the 45 peptides designed and evaluated herein are presented in Table 1. Peptides Sp131–Sp139 were directed for experimental validation of our earlier proteochemometric model for Michaelis (K_m) and cleavage rate constants (k_{cat}) of dengue protease substrates [15]. Sp131–Sp139 were of the same type as in our previous study, Abz-RRRRXXXX-nY-NH_2 [15], and were designed to gain low K_m and/or low k_{cat} values; i.e. to bind tight with low rate of cleavage. These were designed by first selecting amino acids at each position that were predicted to be favorable by the proteochemometric model. Selected amino acids were W, Y, G, or M for the P′1 position; W, Y, G, or V for the P′2 position; C for the P′3 position; and C or G for the P′4 position. This thus gave $4 \times 4 \times 1 \times 2 = 32$ possible peptide sequences. Applying D-optimal design on this set gave a diverse and representative subset, comprising Sp131–Sp139. Experimental testing showed that six of the nine peptides exhibited high inhibitory activity. Only peptides with G at the P′1 and/or P′2 positions were inactive.

Peptides SRP-1–SRP-36 were then designed aiming to achieve higher inhibitory activity, and to study the effects of truncation at the cationic non-prime side. As seen from Table 1, peptides SRP-1–SRP-6, which have the general structure Abz-RRRRHXCX-nY-NH_2, show K_i values in the range of 1–10 μM on all four proteases. Peptides SRP-13 and SRP-17, which are analogs of SRP-3 and SRP-5, which lack the Abz and nY groups, show even higher inhibitory activities. This thus indicates that the fluorogenic groups originally used for assaying substrate cleavage do not contribute to the protease inhibition.

Truncations at the P side generally led to decreased activity (cf. results for SRP-13, SRP-14, SRP-15, and SRP-16). Of the four

Table 1
Inhibition of dengue virus NS2B–NS3 protease by peptide inhibitors (K_i values are in μM).

Name	Sequence	K_i (DEN-1)	K_i (DEN-2)	K_i (DEN-3)	K_i (DEN-4)
Sp131	Abz-RRRRYWCG-nY-NH_2	1.3	2.1	3.0	–[a]
Sp132	Abz-RRRRGWCG-nY-NH_2	No inhib.[b]	No inhib.	No inhib.	–
Sp133	Abz-RRRRWYCG-nY-NH_2	0.7	2.0	1.9	–
Sp134	Abz-RRRRMVCG-nY-NH_2	1.7	3.2	2.1	–
Sp135	Abz-RRRRMGCG-nY-NH_2	No inhib.	No inhib.	No inhib.	–
Sp136	Abz-RRRRWWCG-nY-NH_2	1.3	2.2	2.2	2.9
Sp137	Abz-RRRRMYCC-nY-NH_2	2.5	5.5	4.0	4.1
Sp138	Abz-RRRRYVCC-nY-NH_2	2.3	4.8	4.0	3.1
Sp139	Abz-RRRRGGCC-nY-NH_2	No inhib.	No inhib.	No inhib.	No inhib.
SRP-01	Abz-RRRRHWCC-nY-NH_2	1.3	3.7	1.6	3.0
SRP-02	Abz-RRRRHWCA-nY-NH_2	0.6	1.9	0.9	2.5
SRP-03	Abz-RRRRHWCW-nY-NH_2	0.6	1.1	0.7	3.3
SRP-04	Abz-RRRRHLCC-nY-NH_2	2.8	8.3	5.4	2.9
SRP-05	Abz-RRRRHLCA-nY-NH_2	0.6	3.6	1.4	2.9
SRP-06	Abz-RRRRHLCW-nY-NH_2	0.3	1.1	0.5	1.9
SRP-07	Abz-RRRRWW-nY-NH_2	2.7	44.3	5.3	3.1
SRP-08	Abz-RRRRWY-nY-NH_2	10.3	26.8	9.8	4.7
SRP-09	Abz-RRWWCC-nY-NH_2	1.7	2.7	2.3	4.7
SRP-10	Abz-RRWYCG-nY-NH_2	3.3	4.7	2.8	9.8
SRP-11	Abz-WWCC-nY-NH_2	No inhib.	No inhib.	No inhib.	No inhib.
SRP-12	Abz-WYCG-nY-NH_2	10.2	13.5	20% at 100 μM	20.1
SRP-13	RRRRHWCW-NH_2	0.4	0.3	0.5	2.2
SRP-14	RRHWCW-NH_2	5.4	8.7	5.2	8.5
SRP-15	RHWCW-NH_2	13.1	142.6	25.7	24.9
SRP-16	HWCW-NH_2	No inhib.	No inhib.	No inhib.	No inhib.
SRP-17	RRRRHLCW-NH_2	0.3	0.4	0.7	2.2
SRP-18	RRHLCW-NH_2	5.7	9.9	7.5	9.1
SRP-19	RHLCW-NH_2	No inhib.	No inhib.	No inhib.	No inhib.
SRP-20	HLCW-NH_2	No inhib.	No inhib.	No inhib.	No inhib.
SRP-21	RWYCG-NH_2	No inhib.	No inhib.	No inhib.	No inhib.
SRP-22	WYCG-NH_2	No inhib.	No inhib.	No inhib.	No inhib.
SRP-23	RWWCC-NH_2	19.5	39.0	15.4	9.5
SRP-24	WYCW-NH_2	4.2	4.8	24.4	11.2
SRP-25	Ac-RRRRHWCW-NH_2	0.4	0.3	0.6	2.7
SRP-26	Ac-RRHWCW-NH_2	5.0	8.2	5.2	6.2
SRP-27	Ac-RHWCW-NH_2	16.6	6.3	25.8	17.9
SRP-28	Ac-HWCW-NH_2	No inhib.	No inhib.	No inhib.	No inhib.
SRP-29	Ac-RRRRHLCW-NH_2	0.6	1.1	0.6	2.6
SRP-30	Ac-RRHLCW-NH_2	10.1	5.1	17.0	11.5
SRP-31	Ac-RHLCW-NH_2	No inhib.	No inhib.	No inhib.	No inhib.
SRP-32	Ac-HLCW-NH_2	No inhib.	No inhib.	No inhib.	No inhib.
SRP-33	Ac-RWYCG-NH_2	No inhib.	No inhib.	No inhib.	No inhib.
SRP-34	Ac-WYCG-NH_2	No inhib.	No inhib.	No inhib.	No inhib.
SRP-35	Ac-RWWCC-NH_2	45.1	9.5	37.6	42.8
SRP-36	Ac-WYCW-NH_2	9.0	18.0	75.1	55.9

[a] Not tested.
[b] No activity at 100 μM.

tetrapeptides that lacked entirely arginine residues, three were found to be inactive. However, one with the sequence WYCW-NH_2 (SRP-24) showed inhibitory K_i values ranging from 4.8 to 24.4 μM on the four dengue proteases.

Another interesting observation was that, while tetrapeptide SRP-22 (WYCG-NH_2) was inactive, peptide SRP-12 (Abz-WYCG-nY-NH_2) showed high inhibitory activity against three of the four proteases. We may speculate that this is due to that no steric hindrance takes place from the neighboring G residue so that nY can adopt a conformation that is favorable for interactions with the protease.

Peptides SRP-25–SRP-36 were directed to investigate the influence of acetylation of the N-terminus. These peptides were synthesized in parallel with the non-acetylated analogs (SRP-13–SRP-24). In most cases the acetylated peptides showed similar or up to two-fold lower activity than the non-acetylated ones. Nevertheless, in one case (SRP-27) the activity was higher than for its non-acetylated variant (SRP-15) (6.3 μM versus 142.6 μM against the DEN-2 protease).

We equipped the first 24 peptides of the series (Sp131–Sp139 and SRP-1–SRP-12) with internally quenched fluorogenic groups to get an indication if a peptide was cleaved by the protease and thus acted as a competitive substrate rather than inhibitor. But no increase of background fluorescence during the pre-incubation with the protease was evident for any of these peptides. Peptides SRP-13–SRP-36 were not labeled by fluorogenic groups because they either were shortened or acetylated analogs of SRP-1–SRP-12 or did not contain the cationic substrate cleavage site residues.

3.2. Analysis of the interaction data by proteochemometric modeling

In order to estimate quantitatively the influence of peptide properties for inhibitory activity we performed PCM modeling. To this end, the peptides were characterized by five binary and nine quantitative physico-chemical descriptors while the proteases were described by four binary descriptors (see Section 2.4 for details). The relationship between the descriptors of the 45 peptides and four proteases to the logarithmically transformed inhibitory activity data (pK_i) was established by partial least squares regression, which created a single unified PCM model. The PCM model comprised five PLS components, explaining $R^2 = 0.83$ of the variation of the pK_i data, and having the predictive ability according to fivefold cross-validation $Q^2 = 0.69$.

The correlation of calculated versus measured pK_i values by the PCM model is presented graphically in Fig. 1. As indicated by the ovals in the figure, the activities of three short peptides, namely SRP-12, SRP-24, and SRP-27, were underestimated by about one logarithmic unit, while the activities for one peptide, SRP-36, were overestimated for the four proteases. For most of the peptides calculation errors were below 0.3 logarithmic units, which in fact is comparable to the measurement errors of the inhibition assays.

PLS derives a regression equation where the regression coefficients indicate the influence of the described properties on the response [23]. Since the response in the PCM model was the negative logarithm of the measured K_i values, a positive coefficient indicates that the represented property increases the inhibitory activity, while a negative coefficient indicates the opposite.

As seen in Fig. 2, high positive coefficients are given to the R2 and R4 descriptors, which characterize the presence of at least two (R2) or four (R4) arginine residues in the peptides. A positive coefficient is also given to the Abz-nY descriptor, which represents the presence of fluorogenic groups. When one interprets the coefficient for this descriptor one should take into account that fluorogenic groups are predominantly present in the longest peptides, and hence that the descriptor correlates with the R2 and R4

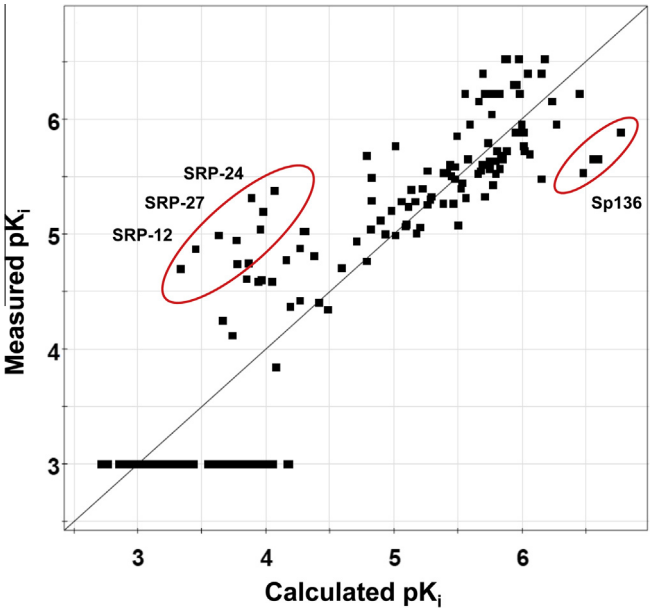

Fig. 1. Correlation of calculated versus measured inhibition activities. Shown are calculated versus measured pK_i values for inhibition of DEN-1–4 virus S2B–NS3 protease by the 45 peptides studied herein, according to the PCM model developed without use of cross-terms (for details see text).

descriptors. Nevertheless, the Abz-nY descriptor explains partially the high activity of the SRP-12 peptide.

Analyzing the *z*-scale descriptors gives further information. As the *z*2-scale represents steric properties of amino acids, the large positive coefficient for 1*z*2 indicates that placing a small amino acid at P′1 should result in peptides with essentially no inhibitory activity. (A plausible explanation for this is that the small G and S, which are the most common amino acids at P′1 among substrates of dengue proteases, facilitate substrate cleavage, as is indicated by our previous PCM modeling of k_{cat} data.) The *z*1-scale characterizes hydrophilicity of amino acids. Hence, the negative coefficients for the 1*z*1 and 4*z*1 descriptors indicate a need for hydrophobic amino acids in the peptide, to achieve inhibitory activity.

The regression coefficients for *z*-scale descriptors at the P′2 position are relatively small and within their confidence intervals, which means that one should not draw too strong conclusions about the preferable amino acid(s) at this position. In fact, this may mean that different amino acids are preferred at this position in different peptides. Detailed inspection of the activity data in Table 1 suggests that a broader variation of amino acids is acceptable in the longest peptides (e.g. aliphatic V and L are tolerated, as well as aromatic W and Y), whereas in the shortest peptides the requirements are more strict (e.g. rendering HWCW-NH_2 and HLCW-NH_2 inactive). To test this hypothesis, we created a new PCM model, which included nine cross-terms obtained by multiplying descriptors 2*z*1, 2*z*2, and 2*z*3 with R1, R2, and R4. The model where 18 original descriptors were complemented with 9 cross-terms showed superior performance; the goodness of fit being $R^2 = 0.92$ and the predictive ability $Q^2 = 0.79$ (compared these values with the $R^2 = 0.83$ and $Q^2 = 0.69$ of the initial model), and the model could better explain the high activities of SRP-12, SRP-24, and SRP-27.

In an attempt to further improve the predictive ability of the PCM model we complemented the 14 descriptors with cross-terms between peptide and protease descriptors. In PCM, use of such cross-terms allows explaining the differences in ligand–protein selectivity. However, few of the cross-terms obtained significant regression coefficients (i.e. they did not fall outside their

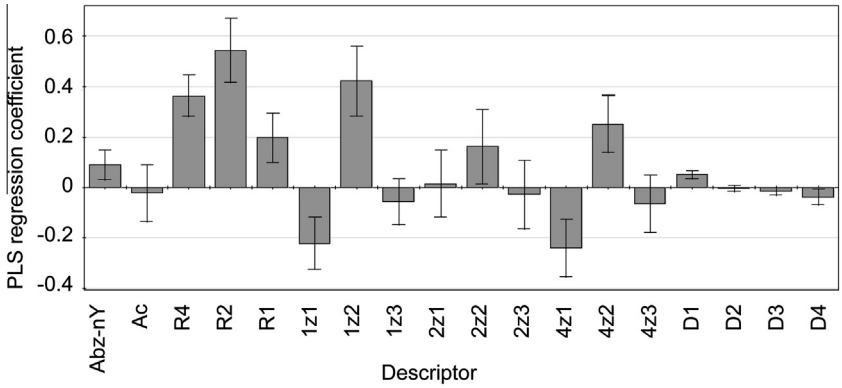

Fig. 2. Influence of peptide and protease properties on the inhibitory activity. Shown are regression coefficients from the PLS regression equation correlating peptide and protease descriptors to their inhibitory activity (pK_i) against proteases of DEN-1–4 virus (i.e. the PCM model without cross-terms). A positive coefficient indicates a positive correlation for the described peptide property and the inhibitory activity, whereas a negative coefficient indicates a negative correlation. Shown are also confidence intervals of regression coefficients calculated from fivefold cross-validation at a confidence level of 95%.

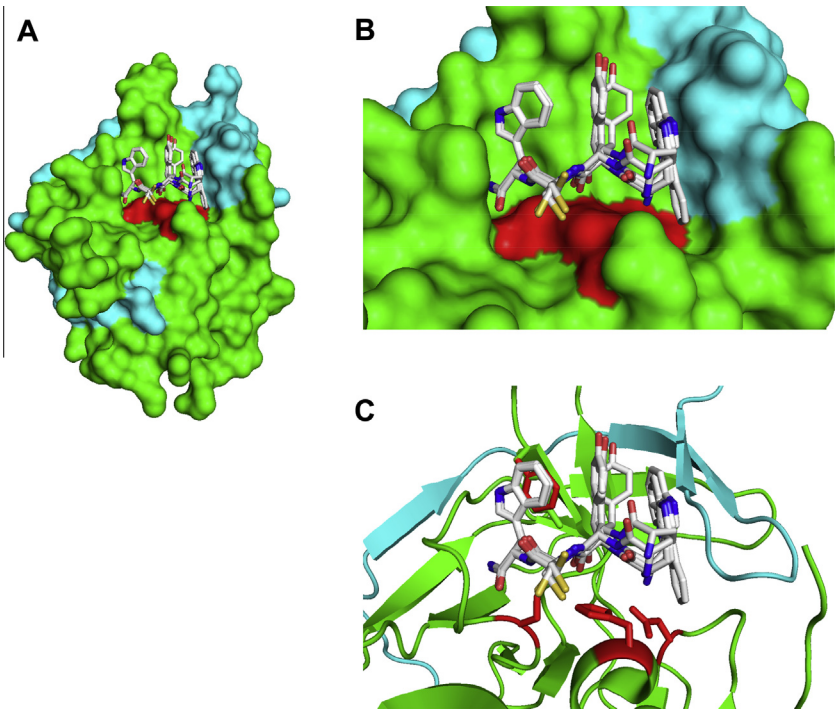

Fig. 3. Results of docking of peptide SRP-24 (WYCW-NH_2) into X-ray structure of NS2B–NS3 protease of DEN-3 virus. (A) Surface view of dengue 3 virus NS2B–NS3 protease (PDB code 3U1I) with docked tetrapeptide WYCW-NH_2. Shown are superimpositions of the lowest energy docking poses of five docking runs. The NS2B unit is shown in cyan, and the NS3 unit in green. The catalytic triad (H51, D75, and S135) is colored red. (B) Close-up of the surface view of the active site with the docked peptide. (C) Close-up of a ribbon view of the active site with docked peptide. Red sticks represent catalytic triad residues H51, D75, and S135, and S1 pocket residue Y161. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

confidence intervals), and the obtained model did not show improved predictive performance. All of this can be explained by the quite low specificity of the studied inhibitors towards some of DEN-1–4 subtypes, something which is in fact highly warranted when one aims to design a broadly acting antiviral.

A drawback of the first derived PCM model, (i.e. the one not exploiting cross-terms), is that it is based on the assumption that all peptides share the same binding mode. While the longest peptides might be 'anchored' by the arginines to the non-prime side binding pockets in the proteases, such anchoring can't be present for the shortest peptides. Comparisons of the calculated pK_i values versus the measured ones revealed that our first PCM model underestimated the activities of SRP-12, SRP-24, SRP-27, and SRP-36. (This is clearly seen from the correlation between the calculated versus measured presented graphically in Fig. 1.) However, PCM modeling has the advantage of being able to explain complex effects and reveal cooperative and incompatible properties by the use of cross-terms (e.g. in previous studies we have demonstrated the use of ligand–protein cross-terms to identify the determinants of ligand–protein specificity [19]). Here, the improved performance of the second model exploiting intra-peptide cross-terms suggests the possibility for different binding locations and/or orientations for the tetra- and pentapeptide inhibitors, compared with the octapeptides.

Although most of the substrates of flaviviral NS2B–NS3 proteases contain dibasic amino acids at the P1 and P2 positions, previous studies have shown that tryptophan and the synthetic amino acids (*p*-Me)Phe and (*p*-guanidinyl)Phe are favored equally

as arginine, and that they are even more favored than lysine at the P1 position for covalently binding peptide aldehyde inhibitors of the DEN-2 protease [24]. X-ray studies have confirmed that an aromatic amino acid has the potential for pi stacking with the Y161 or Y150 residue in the S1 pocket of the protease [2].

3.3. Molecular docking

To test the hypothesis that our tetrapeptide inhibitors might interact differently than the longer arginine containing peptides, we performed docking of SRP-24 (WYCW-NH_2) using the recently solved crystal structure of the DEN-3 virus NS2B–NS3 protease in complex with the covalently bound aldehyde inhibitor Bz-nKRR-H (PDB code 3U1I) [4]. Results for the lowest energy docking solutions are show in Fig. 3. The peptide's N-terminus interacts with aspartic acid D75 of the catalytic triad, while its cysteine residue is located in the area for the scissile bond of protease substrates, where it is turned towards H51 and S135 of the catalytic triad or towards the oxyanion hole of the protease. Residue four of our tetrapeptide WYCW-NH_2 forms pi stacking with Y161 in the S1 pocket. (In the crystal structure, the S1 pocket accommodates the benzoyl group and P1 arginine of the aldehyde inhibitor.) The two first residues of our peptide occupy the S3 and S4 pockets (but in some low energy docking solutions the first W was alternatively located to the S2 pocket, rather than to S3). Superimposition of the results from five docking runs shows good convergence (Fig. 3). However, other orientations, showing within 0.5–1 kcal/mol energy differences from the best docking pose, were also present in the docking solutions. (In fact, this is not surprising considering the similar properties of three of the four residues in the peptide).

3.4. Summing up

We exploited the results of our previous study on the specificity of dengue protease substrates to design a substrate-based library of peptide inhibitors. Most of the designed octapeptides inhibited proteases of all four serotypes of dengue virus with low micromolar or submicromolar activities. To increase drug-likeliness of the peptides we continued the study by stepwise removal of arginines and further methodical variations of amino acids at the four prime side sequence positions. This resulted in an uncharged tetrapeptide, WYCW-NH_2, with low micromolar inhibitory activities against all four proteases.

Continued studies are now being carried out to explore short uncharged peptide inhibitor-NS2B–NS3 protease interaction mechanisms. To this end, we have designed a tetrapeptide library around SRP-24 by systematically modifying all four sequence residues, and assays are now in progress with dengue proteases. In the here reported series Cys residue is invariantly present in all peptides and seems to play key role for protease inhibition in both the octapeptides and SRP-24. The thiol group is, however, susceptible to oxidation and could be a concern for *in vivo* drug design. The ongoing study therefore includes, *inter alia*, investigation of the replacement of Cys by physicochemically similar natural and synthetic amino acids. The results of this study will be reported elsewhere.

Acknowledgments

This work was supported by grants from the Thailand Research Fund (TRF) BRG 4980008 to GK and from the Swedish International Cooperation Development Agency (SIDA) to both JW and GK.

References

[1] M.G. Guzman, S.B. Halstead, H. Artsob, et al., Dengue: a continuing global threat, Nat. Rev. Microbiol. 8 (2010) S7–S16.
[2] P. Erbel, N. Schiering, A. D'Arcy, et al., Structural basis for the activation of flaviviral NS3 proteases from dengue and West Nile virus, Nat. Struct. Mol. Biol. 13 (2006) 372–373.
[3] J. Lescar, D. Luo, T. Xu, et al., Towards the design of antiviral inhibitors against flaviviruses: the case for the multifunctional NS3 protein from Dengue virus as a target, Antiviral. Res. 80 (2008) 94–101.
[4] C.G. Noble, C.C. She, A.T. Chao, P.Y. Shi, Ligand-bound structures of the dengue virus protease reveal the active conformation, J. Virol. 86 (2012) 438–446.
[5] A. Schüller, Z. Yin, C.S. Brian Chia, et al., Tripeptide inhibitors of dengue and West Nile virus NS2B–NS3 protease, Antiviral Res. 92 (2011) 96–101.
[6] C. Nitsche, M.A. Behnam, C. Steuer, C.D. Klein, Retro peptide-hybrids as selective inhibitors of the Dengue virus NS2B–NS3 protease, Antiviral Res. 94 (2012) 72–79.
[7] P. Niyomrattanakit, S. Yahorava, I. Mutule, et al., Probing the substrate specificity of the dengue virus type 2 NS3 serine protease by using internally quenched fluorescent peptides, Biochem. J. 397 (2006) 203–211.
[8] S.M. Tomlinson, S.J. Watowich, Anthracene-based inhibitors of dengue virus NS2B–NS3 protease, Antiviral Res. 89 (2011) 127–135.
[9] C. Bodenreider, D. Beer, T.H. Keller, et al., A fluorescence quenching assay to discriminate between specific and nonspecific inhibitors of dengue virus protease, Anal. Biochem. 395 (2009) 195–204.
[10] T.S. Kiat, R. Pippen, R. Yusof, et al., Inhibitory activity of cyclohexenyl chalcone derivatives and flavonoids of fingerroot, *Boesenbergia rotunda* (L.), towards dengue-2 virus NS3 protease, Bioorg. Med. Chem. Lett. 16 (2006) 3337–3340.
[11] C. Nitsche, C. Steuer, C.D. Klein, Arylcyanoacrylamides as inhibitors of the Dengue and West Nile virus proteases, Bioorg. Med. Chem. 19 (2011) 7318–7337.
[12] S. Aravapalli, H. Lai, T. Teramoto, et al., Inhibitors of Dengue virus and West Nile virus proteases based on the aminobenzamide scaffold, Bioorg. Med. Chem. 20 (2012) 4140–4148.
[13] M. Ezgimen, H. Lai, N.H. Mueller, et al., Characterization of the 8-hydroxyquinoline scaffold for inhibitors of West Nile virus serine protease, Antiviral Res. 94 (2011) 18–24.
[14] K. Löhr, J.E. Knox, W.Y. Phong, et al., Yellow fever virus NS3 protease: peptide-inhibition studies, J. Gen. Virol. 88 (2007) 2223–2227.
[15] P. Prusis, M. Lapins, S. Yahorava, et al., Proteochemometrics analysis of substrate interactions with dengue virus NS3 proteases, Bioorg. Med. Chem. 16 (2008) 9369–9377.
[16] L. Eriksson, E. Johansson, Multivariate design and modeling in QSAR, Chemom. Intell. Lab. 34 (1996) 1–19.
[17] P. Niyomrattanakit, P. Winoyanuwattikun, S. Chanprapaph, et al., Identification of residues in the dengue virus type 2 NS2B cofactor that are critical for NS3 protease activation, Virol. 78 (2004) 13708–13716.
[18] D. Leung, K. Schroder, H. White, et al., Activity of recombinant dengue 2 virus NS3 protease in the presence of a truncated NS2B co-factor, small peptide substrates, and inhibitors, J. Biol. Chem. 276 (2001) 45762–45771.
[19] J.E. Wikberg, O. Spjuth, M. Eklund, M. Lapins, Chemoinformatics taking biology into account: proteochemometrics, in: R. Guha, A. Bender (Eds.), Computational Approaches in Cheminformatics and Bioinformatics, Wiley, Hoboken, 2012, pp. 57–92.
[20] S. Hellberg, M. Sjöström, B. Skagerberg, S. Wold, Peptide quantitative structure-activity relationships, a multivariate approach, J. Med. Chem. 30 (1987) 1126–1135.
[21] G.M. Morris, R. Huey, W. Lindstrom, et al., Autodock4 and AutoDockTools4: automated docking with selective receptor flexibility, J. Comput. Chem. 16 (2009) 2785–2791.
[22] O. Trott, A.J. Olson, AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading, J. Comp. Chem. 3 (2010) 455–461.
[23] S. Wold, M. Sjöström, L. Eriksson, PLS-regression: a basic tool of chemometrics, Chemom. Intell. Lab. 58 (2001) 109–130.
[24] Z. Yin, S.J. Patel, W.L. Wang, et al., Peptide inhibitors of dengue virus NS3 protease. Part 2: SAR study of tetrapeptide aldehyde inhibitors, Bioorg. Med. Chem. Lett. 16 (2006) 40–41.