MedChemComm

CONCISE ARTICLE

View Article Online View Journal | View Issue

Cite this: *Med. Chem. Commun.*, 2013, **4**, 1354

Received 29th July 2013 Accepted 2nd September 2013

DOI: 10.1039/c3md00219e

www.rsc.org/medchemcomm

Synthesis, activity and structure–activity relationship of noroviral protease inhibitors[†]

Lisheng Deng,^a Zana Muhaxhiri,^b Mary K. Estes,^c Timothy Palzkill,^{ac} B. V. Venkataram Prasad^b and Yongcheng Song^{*a}

The protease of norovirus, an important human pathogen, is essential for the viral replication and, therefore, represents a potential drug target. A series of tripeptide-based inhibitors of the protease were designed, synthesized and tested, among which several potent inhibitors were identified with K_i values as low as 75 nM. The structure–activity relationships of these inhibitors are discussed.

Introduction

Norovirus is the leading cause of epidemic acute gastroenteritis in humans. Using PCR, more than 90% of outbreaks of nonbacterial gastroenteritis in the US are caused by noroviruses.1 Similar data are reported from the Netherlands and other countries in Europe.2-4 Norwalk virus (GI.1) in the genogroup I is the prototype of this group of viruses, while genogroup II viruses, such as Houston virus (GII.4), are now more prevalent.⁵ According to a recent CDC (Centers for Disease Control and Prevention) survey, noroviral infections account for 21 million cases of gastroenteritis annually in the US, among which ~ 2 million require outpatient treatment and $\sim 70~000$ hospitalization.6 While illness caused by norovirus is generally short-lived, mild, and self-healing in healthy people, it can be severe and sometimes fatal, especially among young children, the elderly and immunocompromised individuals.7 Each year noroviruses are responsible for the deaths of \sim 218 000 children less than 5 years old in developing countries.8 In developed countries, these viruses have been associated with deaths in the elderly in nursing homes, and are important pathogens causing chronic and sometimes severe illness in intestinal, solid organ and renal transplant patients.9-11 In addition, noroviruses are classified as class B biodefense agents in the US because of their high stability, extremely low infectious dose, and high rate of susceptibility in all age groups of humans, making them a potential threat to food and water supplies or the environment.

To date, there have been no therapeutics or vaccines for the viral infections and there is therefore a pressing need to find effective countermeasures for noroviruses.

Theoretically, the best way to prevent a virus infection is by vaccination. While a noroviral virus-like particle vaccine is being developed, the broad-spectrum efficacy of this against the multiple types and emergence of new types of noroviruses remains to be determined,¹² Even if vaccines are effective, antiviral drugs, would be useful to treat epidemics and chronic infections, such as acyclovir and oseltamivir used to treat herpes and influenza virus infections, respectively. These compounds have a broad antiviral activity because they target an essential viral protein, *e.g.*, protease or neuraminidase, which often exhibits a high homology among the different serotypes of a viral family.

The genome of Norwalk virus (NV) consists of three open reading frames (ORFs). As shown in Fig. 1a and b, ORF1 encodes a 200 kDa polyprotein, which is cleaved by a protease (NVPro) into six non-structural proteins, including NVPro itself.¹³ Because its function is essential for viral replication, NVPro represents a potential drug target for intervention.

Our previous X-ray crystallographic study revealed NVPro is a chymotrypsin-like protease with the residues Cys139, His30 and Glu54 as the catalytic triad.¹⁴ A mechanism for the enzyme was proposed based on the structural information, as shown in Fig. 1c. Cys139 functions as a nucleophile, His30 as a general base, and Glu54 is used to facilitate the correct orientation of the His30 imidazole ring. A series of peptidomimetic inhibitors of NVPro were recently reported, with the best compound in the series having an IC₅₀ value of 870 nM (Fig. 1d).¹⁵ We also determined the X-ray crystal structures of three tripeptide aldehyde inhibitors **1–3** (Fig. 1d) in complex with NVPro.¹⁶ Here, we report the design, synthesis and activity of this series of tripeptide based inhibitors against two noroviral proteases from Norwalk and Houston viruses. Structure–activity relationships (SAR) of these compounds are also discussed.

^aDepartment of Pharmacology, Baylor College of Medicine, 1 Baylor Plaza, Houston, Texas 77030, USA. E-mail: ysong@bcm.edu; Fax: +1 713-798-3145; Tel: +1 713-798-7415

^bVerna and Marrs McLean Department of Biochemistry and Molecular Biology, Baylor College of Medicine, 1 Baylor Plaza, Houston, Texas 77030, USA

^cDepartment of Molecular Virology and Microbiology, Baylor College of Medicine, 1 Baylor Plaza, Houston, Texas 77030, USA

[†] Electronic supplementary information (ESI) available: Detailed experimental section. See DOI: 10.1039/c3md00219e



Fig. 1 (A) The polyprotein encoded by NV ORF1 is cleaved by NVPro into 6 functional proteins; (B) the peptide sequences of the five cleavage sites, with those highlighted in red being used for our inhibitor design; (C) a proposed mechanism of NVPro; (D) structures and activities of NVPro inhibitors.

Results and discussion

Inhibitor design and SAR

Based on the mechanism of NVPro, we first synthesized five tripeptide aldehyde compounds 2, 4–7, as shown in Table 1. This design exploited the P3–P1 sequences of the five cleavage sites for NVPro (Fig. 1b) with the P1 amino acid being modified. The aldehyde functionality is to attract a possible nucleophilic attack by the –SH of Cys139 to form a covalent bond with the enzyme to increase the binding affinity. It is noted that since this covalent bond is reversible with rapid kinetics, these aldehyde compounds behave as the competitive inhibitors of NVPro, similar to those of rhinovirus 3C protease.¹⁷ The Gln side chain was chosen because the negatively charged Glu side chain could impair the cell membrane permeability of these compounds. In addition, two Me groups were used to mask the δ -CONH₂ to avoid its cyclization with the aldehyde to form a six-membered hemiglutaminal species, as shown below in Scheme 1.

Similar inhibitor design has been applied to target the proteases of other RNA viruses such as rhinovirus and piconavirus.^{17–19} These five compounds were tested for their inhibitory activities against recombinant NVPro using our previously reported method¹⁴ and the results are shown in Table 1. All these five initial compounds were found to be NVPro inhibitors with their K_i values ranging from 0.31–8.2 µM. Compound 5 with a P3–P1 sequence of Gln-Leu-Gln(diMe) was found to be the strongest among these five, having a K_i value of 310 nM. Compound 2 with a sequence of Ala-Leu-Gln(diMe) is also a potent inhibitor with a K_i of 590 nM, while compounds 7

(Asn-Phe-Gln(diMe)) and **4** (His-Leu-Gln(diMe)) possess low μ M inhibition (K_i : 1.6 and 1.9 μ M, respectively). Compound **6** having a sequence of Val-Pro-Gln(diMe) exhibits the weakest activity with a K_i value of 8.2 μ M among this series.

Since compounds **2**, **4** and **5** contain a P2 Leu residue, we next synthesized compound **1** with a P2 Phe as well as **8** with a P2 Pro to compare their activities with those of compounds **7** and **5** with the same P2 residue, respectively. Compound **1** (Z-Leu-Phe-Gln(diMe)-CHO) was found to be a potent inhibitor of NVPro with a K_i value of 120 nM (Table 1), being \sim 3–5-fold more active than compounds **5** and **2** with the natural substrate sequences. On the other hand, compound **8** with a P3–P2 sequence of Leu-Pro is a very weak inhibitor ($K_i = 46.1 \mu$ M).

The activity data from this small set of tripeptide aldehyde inhibitors suggest Leu or Phe is a more favorable P2 residue than Pro. For example, 1 is 380-fold more active than 8 with the same P3 and P1 residues, while the potencies of all inhibitors with P2 = Leu or Phe (*i.e.*, 1, 2, 4, 5 and 7) fall within a relatively narrow range of 0.1-1.9 µM. Our crystallographic studies show the P2 Leu or Phe side chain is deeply inserted into the hydrophobic S2 pocket of the protein well defined by I109, Q110, R112, and V114. However, the cyclic structure of Pro could restrict it from having deep contacts with the S2 pocket. Our SAR analysis further confirms that the van der Waals interactions between the inhibitor P2 and the protein S2 significantly affect the binding affinity. The second SAR is that the P3 residue seems to play a less important role, fine-tuning the inhibitory activity. For compounds 5, 2 and 4 with the same P2/P1, changing P3 residue from polar Gln to small/hydrophobic Ala as

Table 1 Structures and inhibitory activity (K_i) of NVPro inhibitors





Scheme 1 Isomerization of glutaminal.

well as larger/bipolar His only results in ~6-fold activity decrease (0.31 \rightarrow 0.59 \rightarrow 1.9 μ M). Similarly, \leq 10× potency variation is found when altering the P3 Leu in compound **1** to the Asn in 7, or the P3 Val in 6 to the Leu in **8**. This observation could be attributed to the less organized S3 pocket of NVPro, which is actually half exposed to the solvent.

Acetyl containing compound **9** was synthesized to find if the Z (benzyloxycarbonyl) protecting group is important for the inhibition. As compared to compound **1** with the same P3–P1 sequence, **9** ($K_i = 39 \mu$ M) exhibits a large activity loss (>300-fold) for NVPro, showing Z contributes considerably to the tight binding of **1**.

We next investigated the SARs with respect to the P1 residue. Since the aldehyde functionality is highly reactive and may be deactivated by cellular or plasma thiols such as glutathione,17 we synthesized compounds **10** and **11** with a methylketone and α , β unsaturated ethyl ester as a nucleophile, respectively. These two functional groups have been successfully used in inhibitor discovery targeting other viral proteases. The P3-P1 sequence of 10 and 11 is the same as that of compound 1. However, ketone 10 only exhibits a very weak inhibitory activity with a K_i value of 150 μ M, being 1200-fold less active than 1 (Table 1). α , β -Unsaturated ester 11 was found to be essentially inactive. Because the α , β -unsaturated ester functionality is particularly attractive as it is chemically stable to non-specific nucleophilic attacks by cellular thiols,18,19 compounds 12 (with an unmasked Gln P1 residue) and 13 (with a β -acetyl diaminopropionic acid as the P1 residue) were also synthesized to test if a different P1 side chain affects the activity. Neither 12 nor 13 exhibits inhibitory activity against NVPro. These results show a highly reactive electrophile such as aldehyde is required to potently inhibit the protease.

Compounds 3, 14-16 were synthesized to optimize the P1 side chain. As compared with compound 1 with the same P3-P2 sequence of Leu-Phe, compound 3 with a reversed amide P1 side chain is $\sim 12 \times$ less active ($K_i = 1.5 \mu$ M). Compound 14 was designed to find if both N-Me of 1 is important for the activity. 14 was, however, found to possess a largely reduced inhibitory activity with a K_i value of 5.1 μ M, being 42-fold less active than 1. Although the crystal structure of NVPro:1 indicates the N-Me trans to the carbonyl O atom in 1 has hydrophobic interactions with Ala160, it seems unlikely that the reduced van der Waals contacts for 14 (without the corresponding N-Me) is mainly responsible for such a large activity loss. More possibly, compound 14 exists as equilibrium between an aldehyde form and a cyclic hemiglutaminal (Scheme 1), with the latter species being inactive against NVPro. Indeed, the ¹H NMR of compound 14 shows the aldehyde form accounts for <10%. Similarly, compound 15 ($K_i = 1.4 \mu M$) with an *N*-cyclopropyl-Gln side chain is also considerably less active than compound 1. As compared to 14, the $\sim 3 \times$ improved inhibition for 15 might be attributed to a

more favorable hydrophobic interaction between the cyclopropyl with the S1 pocket of NVPro. Compound **16**, an analog of **1** containing a pyrrolidine-amide in the P1 position, was found to be the most potent inhibitor of NVPro with a K_i value of 75 nM. The slightly enhanced activity should be due to the additional van der Waals interactions of the 5-membered pyrrolidine ring with Lys162 and Ser163 in the S1 pocket.

Inhibition of Houston virus protease (HOVPro)

Houston virus (HOV) belongs to the genotype II noroviruses.⁵ As shown in Fig. 2a, these two noroviral proteases exhibit a high sequence homology of 66% identity and 85% similarity. Because of its high prevalence in noroviral infections in recent years, we also tested 4 selected inhibitors, **1–3** and **5**, against the protease of HOV (HOVPro) and the results are shown in Table 2. As compared with NVPro, these compounds exhibited reduced inhibitory activities against HOVPro with K_i values of 4.3–14 μ M. Compound **1** with P2-P1 of Phe-Gln(diMe) shows ~60-fold activity reduction, while compounds **2** and 5 with a Leu at the P2 have ~10-fold decreases, being the best HOVPro inhibitors (K_i = 4.6 and 4.3 μ M, respectively). Although having the same P3-P2



Fig. 2 (A) Alignment of NVPro and HOVPro protein sequence. The conserved residues that compose the S2 pockets of the two proteins are labeled with a bracket and a major difference is highlighted red letters; (B) superposition of the crystal structure of NVPro (PDB: 2FYQ) shown in cyan and a homology model of HOVPro structure in gold. The S2 pocket loop (bll–cll) shows a different conformation in the HOV protease, which appears to be due to the Gly115 residue. In the NV protease the residue at the 115 position is His; (C) tables depict the conserved residues that compose the S1, S2 and S4 pockets, in NV and HOV proteases.

Table 2 K_i (μ M) against NVPro and HOVPro

NVPro	HOVPro
0.12	7.5
0.59	4.6
1.5	14
0.31	4.3
	NVPro 0.12 0.59 1.5 0.31

residues as 1, compound 3 with a β -acetyl diaminopropionic acid as the P1 residue also show $\sim 10 \times$ activity loss.

To find out possible reasons for the activity changes, a homology modeling of HOVPro was performed using the program Phyre²⁰ and the result is shown in Fig. 2b, superimposed together with the crystal structure of NVPro. There is only one major conformational difference between the two structures. The S2 loop of HOVPro (104-115) moves more closely towards the protein active site, with the β -sheet kinked at Gly115. However, in NVPro, the side chain of the corresponding residue His115 forms a H-bond with Glu75, which stabilizes the NVPro S2 loop and prevents it from turning inwards (Fig. 2b). The consequence for the closer S2 loop of HOVPro is a smaller S2 pocket, which favours a smaller P2 residue of an inhibitor. This homology model seems to be reasonable, since compounds 2 and 5 with a smaller Leu as a P2 show better inhibitory activities than compound 1 with a bulkier Phe at this position (Table 2). The generally reduced activities of these inhibitors could also be due to the same reason. Gly115 renders the S2 loop of HOVPro more flexible (with a higher overall entropy), which could reduce the binding affinity of an inhibitor. In addition, as shown in Fig. 2c, the S1 pocket of HOVPro contains more polar residues (i.e., Thr135 and 158 vs. Ile135 and A158 in NVPro). Our experimental results ($\sim 10 \times$ activity reduction for 3 vs. >60× for 1) suggest β -acetyl diaminopropionate at the P1 position could be more favourable for HOVPro.

Compound synthesis

The general methods for synthesis of compounds 1–16 are shown in Scheme 2 and detailed synthesis as well as compound characterization are described in ESI.† EDC (N-(3-dimethylaminopropyl)-N'-ethylcarbodiimide hydrochloride) and HOBT (1-hydroxybenzotriazole) were used as a general peptide/amide coupling method. Amino acid methyl ester 19, representing the P1 component of the final products, was synthesized from Z-Glu-OMe (α -Z-protected glutamic acid methyl ester) or α -Z-protected (S)-2,3-diaminopropionic acid methyl ester. Using an EDC/HOBT coupling reaction followed by hydrogenation to remove the Z protection group, tripeptide 21 was prepared from 19, via a dipeptide intermediate 20. The methyl ester of 21 was selectively reduced to a primary alcohol followed by Dess-Martin periodinane oxidation, to give compounds 1-8 and 14-16 having an aldehyde functionality. For synthesizing compounds 11-13, the corresponding aldehyde was treated with the sodium salt of triethyl phosphonoacetate in tetrahydrofuran to convert the aldehyde to α , β -unsaturated ethyl ester. Compound 10 was prepared by treating compound 1 with MeMgBr at -78 °C



followed by Dess–Martin periodinane oxidation. For making acetyl containing compound **9**, Ac-Leu-OH was used to react with compound **20** to obtain an acetyl protected tripeptide **21**, which was reduced and oxidized to give compound **9**.

Conclusion

In conclusion, a series of 16 tripeptide based compounds were designed and synthesized targeting the protease of noroviruses, which is a potential drug target for treating noroviral infections. Compounds 1 and 16 with a general structure of Z-Leu-Phe- $Gln(N,N-X^1,X^2)$ -CHO were found to be the most potent inhibitors of NVPro with K_i values of 120 and 75 nM, respectively. From the activities of this small set of compounds, SAR analysis shows that (1) a potent NVPro inhibitor requires to have an aldehyde group to be the P1 electrophile; (2) N,N-disubstituted Gln is the most favorable P1 side chain; (3) Phe or Leu is favorable for the P2, while Pro at this position is bad for the activity; and (4) the P3 residue seems to play a minor role in tight binding. Selected NVPro inhibitors were tested against the protease of the genotype II Houston virus. Despite a high homology of these two proteases, NVPro inhibitors exhibited a considerably reduced activity against HOVPro. A homology model of HOVPro was generated, which can rationalize the observed activity changes.

Acknowledgements

This work was supported by grants from the NIH (PO1 AI057788 to M.K.E., T.P. and B.V.V.P., and P30DK56338 to M.K.E.), the

Robert Welch Foundation (Q1279 to B.V.V.P.), the Caroline Wiess Law Fund for Molecular Medicine and a seed grant from Texas Medical Center Digestive Disease Center (to Y.S.).

Notes and references

- 1 M. M. Patel, A. J. Hall, J. Vinjé and U. D. Parashar, *J. Clin. Virol.*, 2009, **44**, 1.
- 2 A. D. Hale, T. N. Tanaka, N. Kitamoto, M. Ciarlet, X. Jiang, N. Takeda, D. W. Brown and M. K. Estes, *J. Clin. Microbiol.*, 2000, **38**, 1656.
- 3 B. A. Lopman, D. W. Brown and M. Koopmans, *J. Clin. Virol.*, 2002, 24, 137.
- 4 J. Vinje, S. A. Altena and M. P. Koopmans, *J. Infect. Dis.*, 1997, **176**, 1374.
- 5 K. Y. Green, in *Fields Virology*, ed. D. M. Knipe and Lippincott, Williams & Wilkins, Philadelphia, 2007, pp. 949–980.
- 6 A. J. Hall, B. A. Lopman, D. C. Payne, M. M. Patel,
 P. A. Gastañaduy, J. Vinjé and U. D. Parashar, *Emerging Infect. Dis.*, 2013, 19, 1198.
- 7 K. Bok and K. Y. Green, N. Engl. J. Med., 2012, 367, 2126.
- 8 M. M. Patel, M. A. Widdowson, R. I. Glass, K. Akazawa, J. Vinje and U. D. Parashar, *Emerging Infect. Dis.*, 2008, 14, 1224.
- 9 D. F. Florescu, L. A. Hill, M. A. McCartan and W. Grant, *Pediatr. Transplant.*, 2008, **12**, 372.
- 10 C. I. Gallimore, D. Lewis, C. Taylor, A. Cant, A. Gennery and J. J. Gray, *J. Clin. Virol.*, 2004, **30**, 196.
- S. S. Kaufman, N. K. Chatterjee, M. E. Fuschino, M. S. Magid, R. E. Gordon, D. L. Morse, B. C. Herold, N. S. LeLeiko, A. Tschernia, S. S. Florman, G. E. Gondolesi and T. M. Fishbein, *Am. J. Transplant.*, 2003, 3, 764.

- 12 R. L. Atmar and M. K. Estes, *Expert Rev. Vaccines*, 2012, **11**, 1023.
- 13 M. Asanaka, R. L. Atmar, V. Ruvolo, S. E. Crawford, F. H. Neill and M. K. Estes, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 10327.
- 14 C. E. Zeitler, M. K. Estes and B. V. V. Prasad, *J. Virol.*, 2006, **80**, 5050.
- 15 K. C. Tiew, G. He, S. Aravapalli, S. R. Mandadapu, M. R. Gunnam, K. R. Alliston, G. H. Lushington, Y. Kim, K. O. Chang and W. C. Groutas, *Bioorg. Med. Chem. Lett.*, 2011, 21, 5315.
- 16 Z. Muhaxhiri, L. Deng, S. S. Sreejesh, B. Sankaran, M. K. Estes, T. Palzkill, Y. Song and B. V. V. Prasad, *J. Virol.*, 2013, 87, 4281.
- S. E. Webber, K. Okano, T. L. Little, S. H. Reich, Y. Xin,
 S. A. Fuhrman, D. A. Matthews, R. A. Love,
 T. F. Hendrickson, A. K. Patick, J. W. Meador, 3rd,
 R. A. Ferre, E. L. Brown, C. E. Ford, S. L. Binford and
 S. T. Worland, *J. Med. Chem.*, 1998, 41, 2786.
- 18 P. S. Dragovich, T. J. Prins, R. Zhou, S. E. Webber, J. T. Marakovits, S. A. Fuhrman, A. K. Patick, D. A. Matthews, C. A. Lee, C. E. Ford, B. J. Burke, P. A. Rejto, T. F. Hendrickson, T. Tuntland, E. L. Brown, J. W. Meador, 3rd, R. A. Ferre, J. E. Harr, M. B. Kosa and S. T. Worland, *J. Med. Chem.*, 1999, **42**, 1213.
- P. S. Dragovich, S. E. Webber, R. E. Babine, S. A. Fuhrman, A. K. Patick, D. A. Matthews, S. H. Reich, J. T. Marakovits, T. J. Prins, R. Zhou, J. Tikhe, E. S. Littlefield, T. M. Bleckman, M. B. Wallace, T. L. Little, C. E. Ford, J. W. Meador, 3rd, R. A. Ferre, E. L. Brown, S. L. Binford, D. M. DeLisle and S. T. Worland, *J. Med. Chem.*, 1998, 41, 2819.
- 20 L. A. Kelley and M. J. E. Sternberg, Nat. Protoc., 2009, 4, 363.