Global Health & Medicine. 2022; 4(6):296-300.

DOI: 10.35772/ghm.2022.01066

Drug development targeting SARS-CoV-2 main protease

Haydar Bulut*

The Experimental Retrovirology Section, HIV and AIDS Malignancy Branch, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA.

Abstract: The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and its variants are responsible for the devastating coronavirus disease 2019 (COVID-19) pandemic with more than 6.5 million deaths since 2019. Although a number of vaccines significantly reduced the mortality rate, a large number of the world population is yet being infected with highly contagious omicron variants/subvarints. Additional therapeutic interventions are needed to reduce hospitalization and curb the ongoing pandemic. The activity of the SARS-CoV-2 enzyme; chymotrypsin-like main protease (M^{pro}) is essential for the cleavage of viral nonstructural polypeptides into individual functional proteins and therefore M^{pro} is an attractive drug target. The aim of this review is to summarize recent progress toward the development of therapeutic drugs against M^{pro} protease.

***Keywords*:** COVID-19, SARS-CoV-2, main protease, drug development

Introduction

The ongoing COVID-19 pandemic initially started with SARS-CoV-2 infection in China, and since then it has evolved to more contagious mutant variants such as delta, and omicron (*1*). The original SARS-CoV2 originated in Wuhan, China has disappeared in the meantime. Many vaccines, including the most effective ones based on novel mRNA technology, cannot fully stop infections (*2,3*). In addition, high mutation rates of coronavirus decreased the vaccine efficacy (*4*).

The genome of SARS-CoV-2 consists of single-stranded positive-sense RNA, which encodes two non-structural polyproteins and several structural & accessory proteins (Figure 1). Non-structural proteins are initially produced in two segments, the shorter polypeptide pp1a contains around 11 proteins, and the larger pp1ab consists of 16 different proteins. Those polypeptide chains need to be processed into single functional units to assemble into new viruses (Figure 1). SARS-CoV-2 contains two different enzymes responsible for the proteolysis of the non-structural polypeptides into single functional proteins. While papain-like protease (PL^{pro}) cleaves the polyproteins at three different sites, SARS-CoV-2 chymotrypsin-like main protease (M^{pro}) cleavage reaction takes place at eleven sites. Inhibition of those viral protease enzymes effectively interrupts the formation of functional viral proteins required for the viral life cycle (*1*).

M^{pro} is one of the most heavily studied drug target in terms of therapeutic development for treating COVID-19, and so far more than 2,700 structures of the M^{pro} have been submitted to protein data bank (PDB) mostly in complex with drug candidates & fragments. Actually, drug development targeting M^{pro} started with emerging SARS-CoV-1 and MERS coronavirus infections prior to SARS-CoV-2 (*5*). Some of the lead compounds had already been designed for the M^{pro} of SARS-CoV-1.

Enzymatic activity and Substrate recognition M^{pro}

M^{pro} functions as a dimer, consisting of 3 domains for each protomer. While two catalytic domains consist of beta sheets forming a substrate binding cleft, the C-terminal domain consists of entirely alpha helices that function as dimerization platforms by interacting with N-terminal residues from the second protomer. Mutations or truncations of N-terminal residues result in a dramatic reduction in enzyme activity, based on this observation several attempts have been made to design dimerization inhibitors that mimic the N-terminal residues, however, this approach has shown limited success so far (*6*).

Natural cleavage sites of M^{pro} have been analyzed in molecular detail in recent studies (*7,8*). Except for the fully conserved glutamine residue at P1 position, the eleven cleavage sites show little conservation. Interestingly, the P1' position within the nsp8-nsp9 sequence is uniquely conserved among the various coronaviruses. In contrast to other cleavage sites, which are occupied by rather small residues of Ser or Ala at P1' position, the nsp8-nsp9 sequence is substituted with

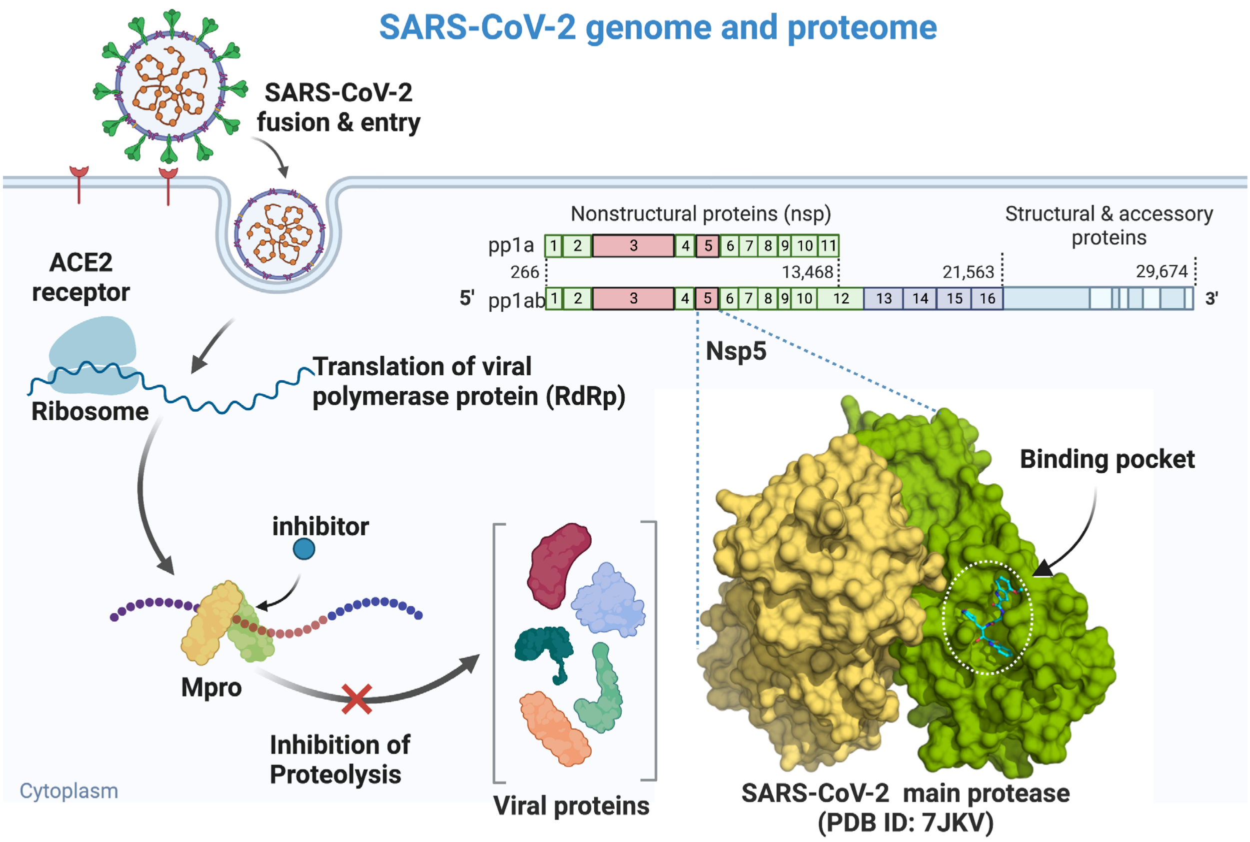
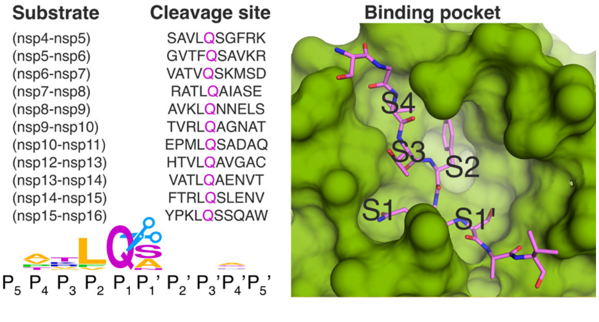


Figure 1. Schematic representation of the SARS-CoV-2 docking onto human ACE2 receptor and entry into the cell. Inside the cell, SARS-CoV-2 replicates its RNA genome in order to manufacture numerous viral proteins using host ribosomes. The proteome of SARS-CoV-2 consists of two non-structural polyproteins pp1a (490 kDa), pp1ab (794 kDa), and additional structural & accessory proteins (pale cyan) such as spike, envelope, membrane, and nucleocapsid. Protomers of M^{pro} is shown in green and yellow, breaking polypeptides into functional viral proteins. Inhibition of M^{pro} with small compound blocks this essential step in viral life cycle. Figure created using Biorender (*https://biorender.com*).

Asn residue at P1' position (Figure 2). Structural and biochemical studies reveal that the presence of Asn of nsp8/9 at P1' position actually reduces the speed of catalytic reaction about 36-fold compared to P1' Ser of nsp4/5 (*9*).

M^{pro} active site contains a cysteine-histidine catalytic dyad (C145 and H41). Key steps of the catalytic cycle are depicted in Figure 3 as the formation of the "thiohemiketal" group, intermediate acyl-enzyme complex, and the final stage cleavage of Gln-(Ser/Ala/Asn) peptide bond with the action of a catalytic water molecule (*10*).



Substrate	Cleavage site
(nsp4-nsp5)	SAVLQSGFRK
(nsp5-nsp6)	GVTFQSAVKR
(nsp6-nsp7)	VATVQSKMSD
(nsp7-nsp8)	RATLQAIASE
(nsp8-nsp9)	AVKLQNNELS
(nsp9-nsp10)	TVRLQAGNAT
(nsp10-nsp11)	EPMLQSADAQ
(nsp12-nsp13)	HTVLQAVGAC
(nsp13-nsp14)	VATLQAENVT
(nsp14-nsp15)	FTRLQSLENV
(nsp15-nsp16)	YPKLQSSQAW

Figure 2. The left 10 amino acid length of M^{pro} cleavage sites within pp1ab is illustrated. The logo was generated using WEBLOGO (*http://weblogo.threeplusone.com*) with the recognition sequences of M^{pro} ranging from the P5 to P5' positions. Schematic diagram of subsite binding pockets with site-specific residues is indicated on the right. While the S1 subsite only recognizes Gln at this position, the S2 site recognizes hydrophobic residues such as L, F or V, and other subsites tolerate more variation in peptide sequences.

Molecular details of M^{pro}-Inhibitor interactions

As the COVID-19 pandemic emerged scientists around the world accelerated drug development efforts. One of the first M^{pro} inhibitors redesigned for M^{pro} was the so-called compound N3, which was based on α, β-unsaturated carbonyl groups that form a covalent bond with a catalytic cysteine residue (Cys145) inside the active site (*11*). The chemical composition of the inhibitor resembles the natural substrate of M^{pro} (Figure 4a, 4b). The γ-lactam ring as P1 moiety engages in bifurcated hydrogen bonds as Gln forms at this position. In addition, a larger lactam ring increased productive var der walls interactions (*7,12,13*). Since then, the γ-lactam ring has

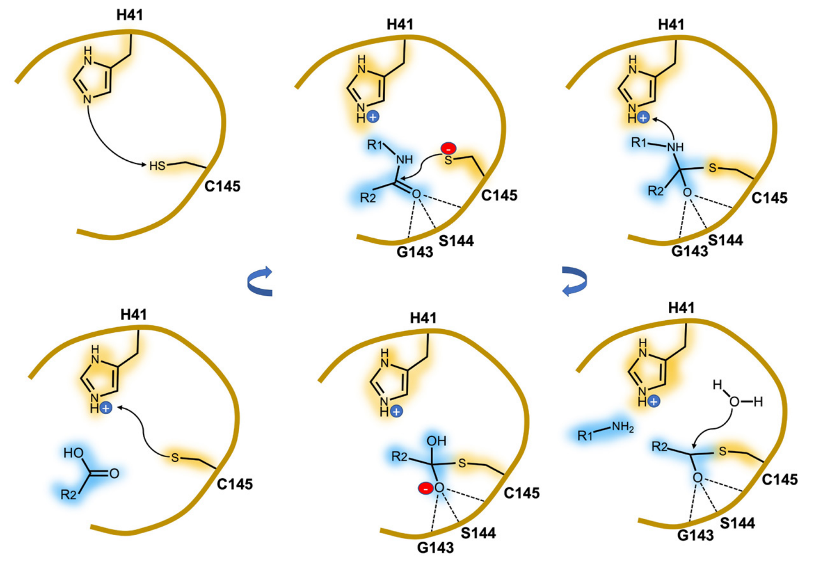


Figure 3. Catalytic mechanism of the proteolysis by M^{pro}. The active site of M^{pro} with the catalytic dyad is highlighted in yellow and the substrate peptide is highlighted in blue. In the free state, imidazole of H41 deprotonates the thiol of C145. The proteolysis reaction starts with the nucleophilic attack by the deprotonated C145 sulfur on the peptide carbonyl carbon. During the transition state, an oxyanion hole forms between the negatively charged oxygen atom and backbone amides of C145, S144, and G143 (dashed triple lines) in order to stabilize the substrate. The part of the peptide (R1-NH) bond breaks down and is released due to the nucleophilic attack by the water molecule onto carboxyl-moiety, and H41 becomes protonated again.

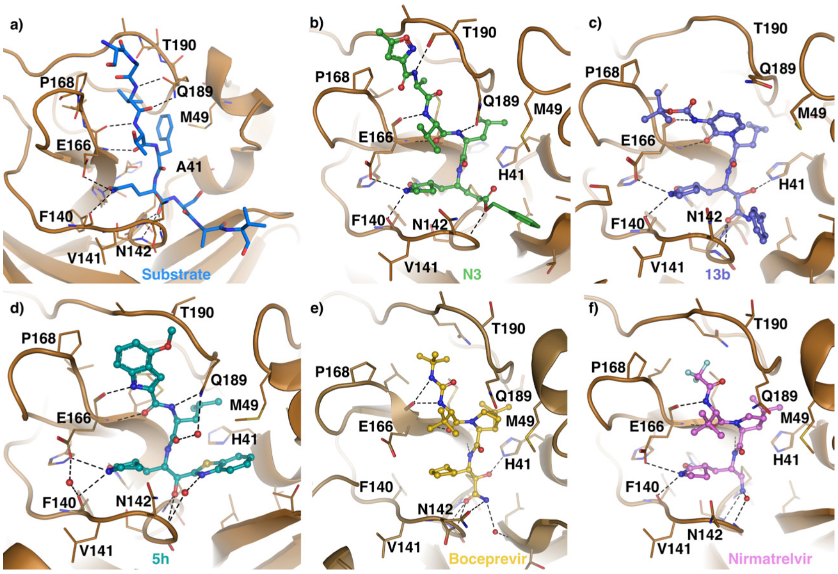


Figure 4. Comparison of the binding modes of substrate and five M^{pro} inhibitors. M^{pros} are shown as cartoons in brown and the inhibitors are shown as stick and ball mode. The binding pocket in complex with nsp5|6 peptidyl substrate (**a**, PDB: 7DVW, blue), N3 (**b**, PDB: 6LU7, green), 13b (**c**, PDB: 6Y2G, purple), 5h (**d**, PDB: 7JKV, aqua-cyan), boceprevir (**e**, PDB: 7BRP, yellow) and nirmatrelvir (**f**, PDB: 7VLQ, pink) are shown. All inhibitors contain P1 gamma lactam in the chemical structure as a common feature except boceprevir. Nirmatrelvir lacks any P1' to occupy S1 subsites.

been essentially kept as a Gln surrogate in the design of most of the M^{pro} inhibitors (*14*). The M^{pro}-N3 complexed structure revealed a strong hydrogen bonding network in a similar fashion to the original substrate, the inhibition of N3 was determined using SARS-CoV-2-infected Vero cells required 10 μM concentration of N3.

Soon later Hilgenfeld's research group designed a new lead compound based on the α-ketoamide group. Initial compound 11r was further optimized by substituting moieties engaging at three subsites resulting in 13b (IC_{50} = 0.67 ± 0.18 μM). The center part of the molecule is built on a pyridine scaffold, which significantly improved drug properties such as plasma half-life and kinetic plasma solubility (Figure 4c). The P2 phenyl group of 11r was substituted with a smaller cyclopropyl methyl moiety, which is deeply embedded in the S2 pocket of M^{pro}. Lead compound 13b was tested on mice with no adverse effects (*15*). Another interesting M^{pro} inhibitor reported 6 months after the pandemic's start was compound 5h comprised of an indole moiety as P4 moiety (*16*). Although the indole group is relatively larger than the substrate peptide at S4 subsite, the M^{pro}-5h-complexed structure revealed a well-fit of the inhibitor inside the binding pocket (Figure 4d). *In vitro* assay performed with VeroE6 cells exposed to SARS-CoVWK-52 resulted in the IC_{50} of 4.2 ± 0.7 μM antiviral activity, which was further boosted with synergistic use of remdesivir (*16*). Recently, 5h was also co-crystallized with MERS and SAR-CoV-1 main proteases revealing similar binding modes with some differences in adaptation of the benzothiazole group inside S1' subsite (*17*). The potency of 5h further increased with the substation of two fluorine atoms (*18*).

So far only one M^{pro}-specific inhibitor "nirmatrelvir" was approved by the FDA to use against COVD-19 (*19*). The oral form of nirmatrelvir/ritonavir is the most effective therapeutic option against SARS-CoV-2 infection reducing hospitalization or death by 89%

(*20*). Nirmatrelvir was developed by Pfizer utilizing the nitrile group as a warhead, which forms covalent bond to the catalytic residue C145. Inside the S2 subsite, the 6,6-dimethyl-3-azabicyclo[3.1.0]hexane group functions well as Leu mimic, and the trifluoro acetyl group inside the S4 subsite engaged in multiple fluorine-based halogen interactions (Figure 4f). Interestingly, nirmatrelvir resembles similar chemical features of hepatitis C virus (HCV) protease inhibitor boceprevir (*21*). Studies show that boceprevir also binds the M^{pro} substrate pocket in a similar conformation (Figure 4e).

Conclusions and future directions

Coronaviruses mutate randomly and active site residues that contact the inhibitor can mutate without affecting the substrate recognition to confer resistance. Inhibitors that optimally occupy the substrate envelope as the natural substrates are less likely to be affected by those mutations. The inhibitors we presented in this review form several hydrogen bonds with protein backbone atoms including oxyanion hole residues through backbone amides, these interactions are likely retained in the active sites of mutant proteases.

Recent studies reported a combination of mutations (L50F and E166V) in the M^{pro} sequence reduces the potency of nirmatrelvir about 80-fold (*22*). Since the P1-γ-lactam ring forms hydrogen bonds with His-163 and Asp-166 side chains, mutation at this P1 site dramatically reduces the inhibitor potency. This possibility of those mutations may emerge in infected people is raising concerns, it may practically end the use of nirmatrelvir.

Previously protease inhibitors have been successfully employed against HIV and HCV proteases and other viral enzymes (*23*). So far about ten HIV protease inhibitors are approved and most of the regiments are given in cocktails to avoid the emergence of new resistant viruses (*23*). The success of anti-HIV-1 protease therapy took a continuous improvement in the potency of inhibitors that are classified in three generations, suggests that anti-M^{pro} drug development is at the initial stage and more potent inhibitors are to arrive, however, more precise understanding of the mechanism of SARS-CoV-2 resistance to M^{pro} inhibitors is required.

Funding: This work was supported by funding from the Intramural Research Program, National Institutes of Health, National Cancer Institute, Center for Cancer Research.

Conflict of Interest: The author has no conflicts of interest to disclose.

Note: The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

References

1. Mistry P, Barmania F, Mellet J, Peta K, Strydom A, Viljoen IM, James W, Gordon S, Pepper MS. SARS-CoV-2 Variants, Vaccines, and Host Immunity. Front Immunol. 2022; 12:809244.
2. Jafari A, Danesh Pouya F, Niknam Z, Abdollahpour-Alitappeh M, Rezaei-Tavirani M, Rasmi Y. Current advances and challenges in COVID-19 vaccine development: from conventional vaccines to next-generation vaccine platforms. Mol Biol Rep. 2022; 49:4943-4957.
3. Pulliam JRC, van Schalkwyk C, Govender N, von Gottberg A, Cohen C, Groome MJ, Dushoff J, Mlisana K, Moultrie H. Increased risk of SARS-CoV-2 reinfection associated with emergence of Omicron in South Africa. Science. 2022; 376:eabn4947.
4. Fan Y, Li X, Zhang L, Wan S, Zhang L, Zhou F. SARS-CoV-2 Omicron variant: recent progress and future perspectives. Signal Transduct Target Ther. 2022; 7:141.
5. Liu Y, Liang C, Xin L, Ren X, Tian L, Ju X, Li H, Wang Y, Zhao Q, Liu H, Cao W, Xie X, Zhang D, Wang Y, Jian Y. The development of Coronavirus 3C-Like protease (3CLpro) inhibitors from 2010 to 2020. Eur J Med Chem. 2020; 206:112711.
6. Goyal B, Goyal D. Targeting the Dimerization of the Main Protease of Coronaviruses: A Potential Broad-Spectrum Therapeutic Strategy. ACS Comb Sci. 2020; 22:297-305.
7. Zhao Y, Zhu Y, Liu X, *et al.* Structural basis for replicase polyprotein cleavage and substrate specificity of main protease from SARS-CoV-2. Proc Natl Acad Sci U S A. 2022; 119:e2117142119.
8. Shaqra AM, Zvornicanin SN, Huang QYJ, Lockbaum GJ, Knapp M, Tandeske L, Bakan DT, Flynn J, Bolon DNA, Moquin S, Dovala D, Kurt Yilmaz N, Schiffer CA. Defining the substrate envelope of SARS-CoV-2 main protease to predict and avoid drug resistance. Nat Commun. 2022; 13:3556.
9. MacDonald EA, Frey G, Namchuk MN, Harrison SC, Hinshaw SM, Windsor IW. Recognition of Divergent Viral Substrates by the SARS-CoV-2 Main Protease. ACS Infect Dis. 2021; 7:2591-2595.
10. Swiderek, K. and V. Moliner, Revealing the molecular mechanisms of proteolysis of SARS-CoV-2 M^{pro} by QM/MM computational methods. Chem Sci, 2020. 11:10626-10630.
11. Yang J, Lin X, Xing N, Zhang Z, Zhang H, Wu H, Xue W. Structure-Based Discovery of Novel Nonpeptide Inhibitors Targeting SARS-CoV-2 M^{pro}. J Chem Inf Model. 2021; 61:3917-3926.
12. Jain RP, Pettersson HI, Zhang J, Aull KD, Fortin PD, Huitema C, Eltis LD, Parrish JC, James MN, Wishart DS, Vederas JC. Synthesis and evaluation of keto-glutamine analogues as potent inhibitors of severe acute respiratory syndrome 3CLpro. J Med Chem. 2004; 47:6113-6116.
13. Dragovich PS, Prins TJ, Zhou R, Webber SE, Marakovits JT, Fuhrman SA, Patick AK, Matthews DA, Lee CA, Ford CE, Burke BJ, Rejto PA, Hendrickson TF, Tuntland T, Brown EL, Meador JW 3rd, Ferre RA, Harr JE, Kosa MB, Worland ST. Structure-based design, synthesis, and biological evaluation of irreversible human rhinovirus 3C

protease inhibitors. 4. Incorporation of P1 lactam moieties as L-glutamine replacements. J Med Chem. 1999; 42:1213-1224.

14. Bai B, Belovodskiy A, Hena M, *et al*. Peptidomimetic α-Acyloxymethylketone Warheads with Six-Membered Lactam P1 Glutamine Mimic: SARS-CoV-2 3CL Protease Inhibition, Coronavirus Antiviral Activity, and *in Vitro* Biological Stability. J Med Chem. 2022; 65:2905-2925.
15. Zhang L, Lin D, Sun X, Curth U, Drosten C, Sauerhering L, Becker S, Rox K, Hilgenfeld R. Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved α-ketoamide inhibitors. Science. 2020; 368:409-412.
16. Hattori SI, Higashi-Kuwata N, Hayashi H, e*t al*. A small molecule compound with an indole moiety inhibits the main protease of SARS-CoV-2 and blocks virus replication. Nat Commun. 2021; 12:668.
17. Hu X, Lin C, Xu Q, Zhou X, Zeng P, McCormick PJ, Jiang H, Li J, Zhang J. Structural Basis for the Inhibition of Coronaviral Main Proteases by a Benzothiazole-Based Inhibitor. Viruses. 2022; 14:2075.
18. Tsuji K, Ishii T, Kobayakawa T, *et al*. Potent and biostable inhibitors of the main protease of SARS-CoV-2. iScience. 2022; 25:105365.
19. Owen DR, Allerton CMN, Anderson AS, *et al*. An oral SARS-CoV-2 M(pro) inhibitor clinical candidate for the treatment of COVID-19. Science. 2021; 374:1586-1593.
20. Hammond J, Leister-Tebbe H, Gardner A, Abreu P, Bao W, Wisemandle W, Baniecki M, Hendrick VM, Damle B, Simón-Campos A, Pypstra R, Rusnak JM; EPIC-HR Investigators. Oral Nirmatrelvir for High-Risk, Nonhospitalized Adults with Covid-19. N Engl J Med. 2022; 386:1397-1408.
21. Fu L, Ye F, Feng Y, *et al*. Both Boceprevir and GC376 efficaciously inhibit SARS-CoV-2 by targeting its main protease. Nat Commun. 2020; 11:4417.
22. Zhou YY, Gammeltoft KB, Ryberg LA, *et al*. Nirmatrelvir Resistant SARS-CoV-2 Variants with High Fitness *in Vitro*. bioRxiv, 2022; doi: https://doi.org/10.1101/2022.06.06.494921
23. Ghosh AK, Osswald HL, Prato G. Recent Progress in the Development of HIV-1 Protease Inhibitors for the Treatment of HIV/AIDS. J Med Chem. 2016; 59:5172-208.

Received November 20, 2022; Revised December 5, 2022; Accepted December 7, 2022.

**Address correspondence to*:
Haydar Bulut, The Experimental Retrovirology Section, HIV and AIDS Malignancy Branch, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA.
E-mail: haydar.bulut@nih.gov