

Structural Evaluations of SARS-CoV-2 Main Protease (M^{pro}): A Review for COVID-19 Antivirals Development Strategy

Muhammad Hamzah Syaifullah Azmi^{1,2}, Ernawati Arifin Giri-Rachman^{1*}

1) School of Life Sciences and Technology, Institut Teknologi Bandung

2) Department of Biotechnology, Wageningen University and Research

*) Corresponding author; e-mail: : erna_girirachman@itb.ac.id

Received: 2023-09-26

Accepted for publication: 2023-12-21

Abstract

It has been almost four years since the first case of COVID-19 emerged, and the antivirals that could work specifically against SARS-CoV-2 with a high efficacy are still under development. Main Protease (M^{pro}) of this virus plays a crucial role in virion maturation during its replication within the host cell. This protein works together with the papain-like protease (PL^{pro}) to cleave polyprotein 1_a and 1_{ab} into a total of 16 functional fragments of non-structural protein. Antiviral with the ability to inhibit the activity of M^{pro} could potentially prevent the virion replication, and they can be developed to target the catalytic or allosteric site of this protein. Antiviral that works on the catalytic site will act as competitive inhibitors of the substrate peptide which leads to the loss of M^{pro} function. Targeting the allosteric site (e.g. distal site and dimerization interface) will cause allosteric modulation of the protomer which could alter the protein 3D conformation and disrupt the formation of homodimer structure. This will affect the geometry and surface structure of the catalytic site which in turn decreases the affinity of the substrate peptide towards the M^{pro} catalytic site, resulting in a complete inactivation of the protein. Mutation study of M^{pro} amino acids sequence also reveals that the mutation frequency for each amino acid position is extremely low and negligible. Moreover, it is found that this protein has 24 mutational cold spot residues scattered within its structure which could be targeted for the development of antivirals due to its highly conserved nature.

Keywords: Antivirus, target sites, mutational cold spot, main protease, SARS-CoV-2

1. Introduction

COVID-19 is known as one of the respiratory tract diseases or acute respiratory syndrome, which was declared a pandemic on March 11, 2020, since it had spread into almost all countries in the world. This pandemic has lasted for almost four years since its first emergence on December 19, 2019, in Wuhan, China [1, 2, 3]. Until now, many countries in the world are still struggling to deal with and prevent the spread of the COVID-19 pandemic. According to global data on September 23, 2023, there have been recorded 695,684,448 cases of infection and 6,919,216 deaths in total due to the pandemic [4, 5]. These numbers directly affect the dynamics of the number of active cases globally. The world has gone through five waves of COVID-19 infection, and it is predicted that in the future, there will be other waves of infection [6, 7]. This is mainly caused by the genome of the COVID-19 pathogen

which has a very high mutation rate. These mutations could promote the occurrence of evolutionary processes leading to the development of various variants of the pathogen. This problem could further make the eradication of the pathogen could not be done easily and cause the emergence of new infection waves [8, 9].

Molecular phylogenetic analysis of the genome of the COVID-19 pathogen reveals that the pathogen originated from the β -Coronavirus (β -CoV) genus which was later named as novel Coronavirus 2019 (nCoV-19) [8, 10, 11]. The name was later changed to Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) on February 11, 2020 [12, 13]. Phylogenetically, this virus belongs to the same genus as SARS-CoV and MERS-CoV, which caused the SARS epidemic in 2002 and the MERS epidemic in 2012 respectively [10, 14]. When these two epidemics are compared with COVID-19, it

can be seen that COVID-19 possesses a much higher transmission or infection rate with a lower mortality rate [15, 16].

Details of SARS-CoV-2 phylogenetic classification can be observed in Table 1.

Table 1. SARS-CoV-2 phylogenetic classification [10]

Realm	<i>Riboviria</i>
Kingdom	<i>Orthornavirae</i>
Phylum	<i>Pisuviricota</i>
Class	<i>Pisoniviricetes</i>
Order	<i>Nidovirales</i>
Family	<i>Coronaviridae</i>
Genera	<i>Betacoronavirus</i>
Subgenera	<i>Sarbecovirus</i>
Species	SARS-CoV-2

As a member of the *Coronaviridae* family, SARS-CoV-2 is classified as a Class IV virus since it carries genetic material in the form of positive single-stranded RNA (ssRNA(+)), sized around 29,000 bps [12, 13]. The characteristic of this kind of genetic material tends to resemble the characteristic of mRNA found in the cytoplasm of the host cell so that the genome of SARS-CoV-2 can be used directly by the host cell's ribosomes as codons for synthesizing the viral proteins [17, 18]. Based on molecular phylogenetic analysis, it is known that the genome of SARS-CoV-2 has a sequence similarity of 79% to SARS-CoV and 51.8% to MERS-CoV [11, 19]. However, when the genetic material of SARS-CoV-2 is compared to the genetic material of CoV that infects bats (bat-CoV RaTG13), a sequence similarity of 96.2% was found. This provides information regarding the origin of SARS-CoV-2 which caused the COVID-19 pandemic [11, 14].

One of the characteristics of RNA viruses, such as CoV, is that they are prone to evolution due to the extremely high mutation rate of their genetic material during the replication process inside the host cell. It is well known that RNA viruses could undergo 10⁻⁶ to 10⁻⁴ substitutions per nucleotide per cell infection [18, 20]. This could be caused by the type of replication enzyme that is used by the virion to replicate its genetic material, which is RNA polymerase. RNA polymerase is further divided into two types based on the template used. If the enzyme uses DNA as its template, the enzyme is known as DNA-dependent RNA polymerase (DdRp or RNAP) and if the template is RNA, the enzyme is known as RNA-dependent RNA polymerase (RdRp) [18, 21, 22]. SARS-CoV-2 which carries genetic material in the form of ssRNA(+) uses RdRp as the key enzyme for the replication of its genetic materials [23]. Unfortunately, RNA polymerase is different from DNA polymerase because this enzyme works by replicating genetic material without proofreading activity. It causes RNA polymerase cannot to repair the errors in the incorporation of the nucleotides during genetic material replication, so RNA polymerase tends to be error-prone [18, 21, 22].

These features of SARS-CoV-2 certainly become the great challenges in developing COVID-19 therapeutic agents,

whether in the form of vaccines or antivirals [24, 25]. Until now, COVID-19 vaccines have been developed and mass-produced on various platforms to build herd immunity within the human population [26, 27]. However, antivirals that can treat COVID-19 infection are still in the stage of development. An ideal antiviral must meet several criteria, including high interaction affinity to the target viral protein (specific interaction), high bioavailability, low manufacturing costs, and low risk of side effects [28, 29]. To date, the only antiviral for the treatment of COVID-19 that has been approved by the Food and Drug Administration (FDA) is remdesivir. In broad terms, remdesivir is a nucleotide-analog compound capable of disrupting the functionality of the RNA-dependent RNA polymerase (RdRp) during the viral RNA replication process. This perturbation has the potential to result in the deactivation of RdRp, thereby impeding the replication of the viral genetic material [30, 31].

In addition to RdRp, SARS-CoV-2 has the main protease (M^{pro}) which plays a crucial role in virion replication and maturation inside the host cell. Because of its importance for the virion, this protein could be a target in the development and design of an antiviral that works as an inhibitor of M^{pro}. Inhibiting the proteolytic activity of M^{pro} will directly suppress the maturation of the SARS-CoV-2 virion which later leads to the inactivity of viral replication [32, 33, 34]. Moreover, SARS-CoV-2 M^{pro} homologs are not present in human cells and its substrate specificity also does not overlap with any known human protease, reducing the probability for the emergence of undesirable effects due to cross reaction by any SARS-CoV-2 M^{pro} inhibitors. These features give more advantage for the development of SARS-CoV-2 antiviral agents targeting this protein specifically [35, 36].

From a structural perspective, It is imperative to address certain crucial factors when endeavoring to design inhibitors targeting the SARS-CoV-2 M^{pro}. The first one is the size and geometry evaluation of the ligand-binding pocket present on the surface of M^{pro} [37]. As a protein, the folding of M^{pro} produces some pockets on its surface, which can be targeted by the inhibitor. Some of these pockets are sufficiently large to

bind a ligand with high stability, while some others are too small or too large that the ligand cannot bind properly within the pocket [38]. Moreover, a ligand should possess the size and structural geometry that matches the size and geometry of the ligand binding pocket of the protein, or a structure that can induce conformational alteration of the ligand binding pocket, improving the geometric fit between the ligand and the pocket [38,39]. The second factor that should be taken into account is the amino acid residues of the protein forming the surface of the pocket. Understanding the essential amino acid residues that require inhibition, along with the chemical characteristics inherent to these residues, holds significance in the process of formulating and enhancing the structural attributes and functional groups within the inhibitor molecule. A desired inhibitor should be able to occupy the pocket and bind to the key amino acids with high stability, inhibiting or altering the conformation of the M^{pro} [40,41]. Consideration of these two

aspects is important during the inhibitor development to increase the efficacy of the inhibitor, thus reducing the concentration needed to effectively inhibit the M^{pro} (reducing the IC_{50} and EC_{50} of the drug). Moreover, higher binding affinity of the ligand to the pocket can also be achieved, which could prolong the effect of the inhibitor to the M^{pro} , increasing its working efficiency [37,40,41].

In order to develop M^{pro} inhibitors that meet the criteria of ideal antiviral agents, detailed studies of M^{pro} , especially its structure, must be carried out. This review was conducted to summarize essential information related to M^{pro} including the aspect of sequence conservation and druggability potential of each part within the M^{pro} structure in a comprehensive, detailed, and holistic manner, so that this review can be used as a reference for the development of antivirals for the treatment of diseases caused by β -CoV in the future.

2. Main Protease (M^{pro}) of SARS-CoV-2

The genetic material of SARS-CoV-2 consists of 14 open reading frames (ORFs). The biggest ORF that can be found within the genetic material is the $ORF1_{ab}$, which is sized for about 65% of the total size of SARS-CoV-2 genome. $ORF1_{ab}$ carries the coding sequence of both polyprotein 1_a ($pp1_a$) and

polyprotein 1_{ab} ($pp1_{ab}$) [13, 42]. The structure of the SARS-CoV-2 genome can be seen in Figure 1. These two polyproteins carry a total of 16 nonstructural proteins (nsps) of SARS-CoV-2 which play an important role during the process of virion replication inside the host cell [42, 43].

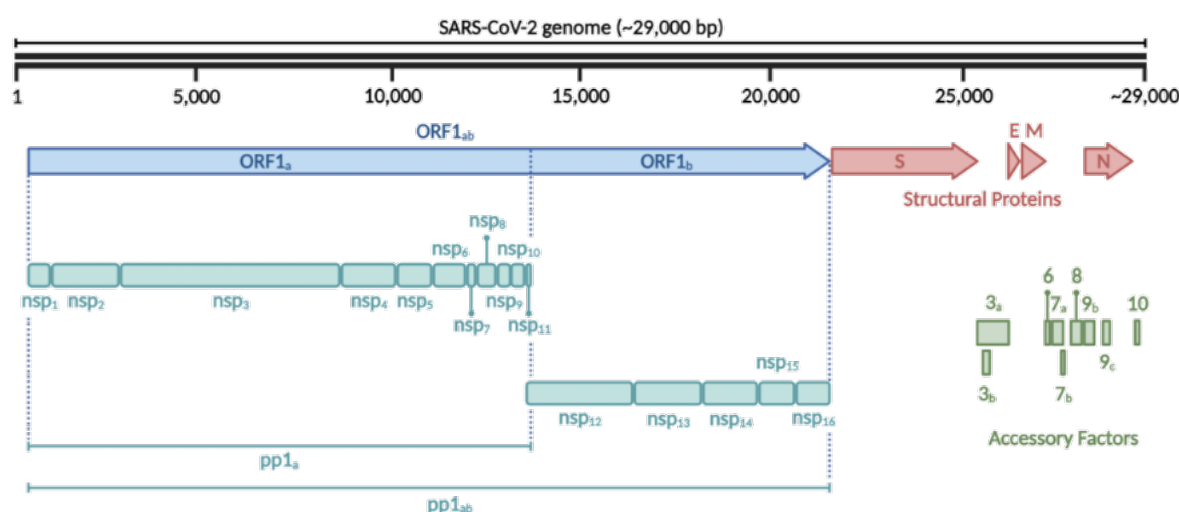


Figure 1. The structure of the SARS-CoV-2 genome. (blue) $ORF1_{ab}$ brings the coding sequence of $pp1_a$ and $pp1_{ab}$ (cyan), containing all 16 SARS-CoV-2 nonstructural proteins. (red) $ORF3$ brings the coding sequence of SARS-CoV-2 structural proteins. (green) $ORF3$ codes for accessory proteins. Figure reproduced and modified from [32] by open access Creative Commons CC BY license. Created with BioRender.com

One of the nsps that can be found in $pp1_a$ and $pp1_{ab}$ is nsp5 which is also known as M^{pro} or 3-chymotrypsin-like protease ($3CL^{\text{pro}}$). The fragment of M^{pro} spans about 306 amino acid residues in both polyproteins from the amino acid position of 3,264th to the 3,569th [33, 44]. This protein, together with nsp3 (papain-like protease, PL^{pro}), is responsible for virion maturation

by cleaving $pp1_a$ and $pp1_{ab}$ into 16 functional nsp fragments. M^{pro} plays a role in catalyzing the proteolysis reaction of seven cleavage sites on $pp1_a$ and eleven cleavage sites on $pp1_{ab}$ [33, 44]. The cleavage sites within $pp1_a$ and $pp1_{ab}$ can be seen in general in Figure 2.

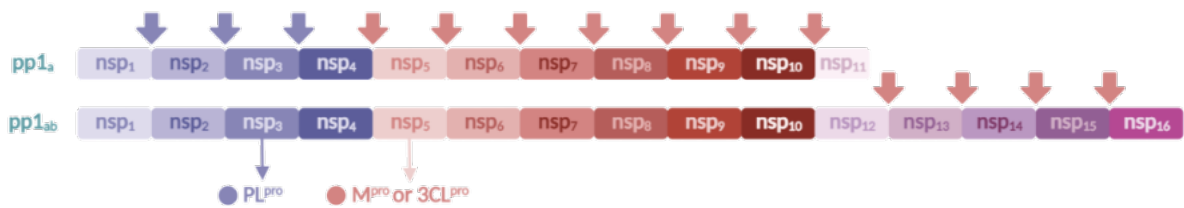


Figure 2. Cleavage sites within the structure of pp1_a and pp1_{ab}. Cleavage sites of pp1_a and pp1_{ab} that can be recognized by PL^{pro} (purple arrow) and M^{pro} (red arrow). Figure reproduced and modified from [33] by open access Creative Common CC BY license. Created with BioRender.com

Crystal structure evaluation of SARS-CoV-2 M^{pro} using the X-ray diffraction method reveals that the native structure of this protein is in a symmetrical homodimer state composed of two identical protomers that interact with each other in a parallel manner [45, 46]. The interactions between the two protomers are formed within the dimerization interface and involve weak intermolecular bindings that occur between the amino acid residues of one protomer and the amino acid residues of the counterpart protomer [33, 45, 46]. Each protomer is sized for about 306 residues of amino acids and has three main domains, namely domain I, domain II, and domain III. In addition to the three main domains, each protomer also has some parts including N-finger, loop, and C-terminal sequence [33, 46]. The details of each domain and part along with their functions can be seen in Table 2, while the structure of SARS-CoV-2 homodimeric M^{pro} can be observed in Figure 3.

Table 2. Structures of SARS-CoV-2 Mpro and its functions [33].

Structures	Range	Length (AA*)	Function
N-finger	Ser1 – Ser10	10	• Dimerization • Increasing the stability of subsite-1 within the substrate-binding pocket of the counterpart protomer
Domain I	Gly11 – Pro99	88	• Together with domain II forms the substrate-binding pocket
Domain II	Lys100 – Tyr182	83	• Together with domain I forms the substrate-binding pocket
Loop	Gly183 – Asp197	15	• Connector of domain II and domain III • Responsible for the protomer structure flexibility
Domain III	Thr198 – Val303	107	• Main part of dimerization • Increasing the intermolecular stability between the two protomers within the homodimer structure
C-terminal	Thr304 – Gln306	3	• Dimerization

*Amino acid residues.

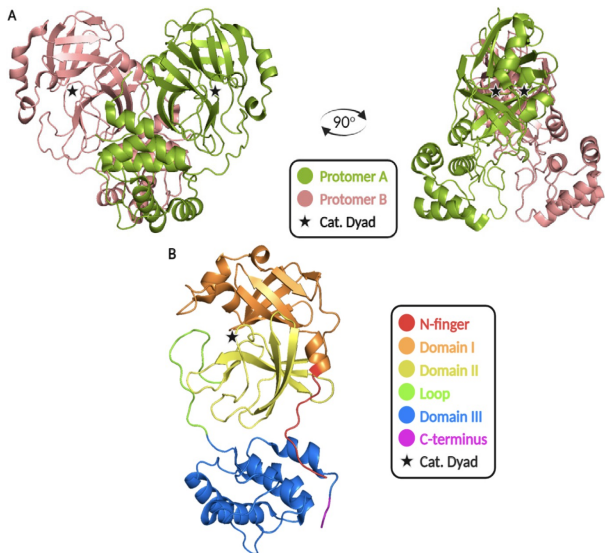


Figure 3. Cartoon model of SARS-CoV-2 M^{pro} structure. (A) The native structure SARS-CoV-2 M^{pro} that is consisted of two identical protomers, the protomer A (green) and the protomer B (pink), the catalytic dyad of each protomer is labeled by star. (B) The structure of each protomer of SARS-CoV-2 M^{pro} in which each structure is labeled by different color. Note that the main domain of the protomer are the domain I, domain II, and domain III. Figure reproduced and modified from [33] by open access Creative Common CC BY license. Created with BioRender.com and PyMOL.

The development of SARS-CoV-2 M^{pro} inhibitors can be carried out by targeting the catalytic site or allosteric site of each protomer. Considering that the geometry of M^{pro} allows this protein to form many ligand-binding pockets scattered all over its surface. The largest and most potentially used ligand-binding pocket as a target is the catalytic site or substrate-binding site [34, 47, 48]. Inhibitors that work on the catalytic site will act as competitive inhibitors against the peptide substrates of M^{pro}. Meanwhile, the ligand-binding pockets apart from the catalytic site is known as allosteric sites, including the dimerization sites and distal sites [34, 49]. Inhibitors acting on the dimerization site can prevent the formation of

the homodimeric structure of M^{pro}. The inhibition of dimerization could further prevent the activation of both protomers so that the enzyme cannot undergo the proteolytic reaction [34, 50]. Moreover, inhibitors acting on the distal site could cause an overall conformational alteration (allosteric modulation) of the M^{pro} protomer. This could lead to a change in space on the catalytic site which can reduce the affinity of the enzyme toward the substrate peptide. The alteration could also cause a shift in the dimerization interface which can reduce the affinity between the two protomers to form a homodimeric state [34, 51, 52].

3. Catalytic Site of SARS-CoV-2 M^{pro}

M^{pro} from the *Coronaviridae* family, including SARS-CoV-2, is a type of cysteine protease. This is because the substrate-binding pocket of each M^{pro} protomer has a couple of catalytic residues, called catalytic dyads, which consists of cysteine (Cys, C) and histidine (His, H) [33, 44]. The substrate-binding pocket or catalytic site of each M^{pro} protomer lies between domain I and domain II with catalytic dyads of His41, located in domain I, and Cys145, located in domain II [53, 54]. The key residues that form the substrate-binding pocket are arranged in such a way as to form the geometry of surface, cavity, and chemical environment that is ideal for the specific geometry of residues of the substrate peptide. The

cavity in the substrate-binding pocket is divided into five subsites or subpockets, namely S1, S2, S3, S4, and S5 [53, 55]. Each subsite is involved in interactions with specific amino acid residues within the substrate peptide. The position of the amino acid residue of the substrate peptide that is bound to the subsite is labeled as P. The proteolytic reaction occurs between the position P1 and P1' [55]. Therefore, the P1, P2, and P1' position greatly determine the specificity of interaction with the corresponding subsite located on the catalytic site of M^{pro}. Meanwhile, the other positions help to increase the stability of interaction and promote the recognition of the substrate by the catalytic site [55]. The structure of the catalytic site along with its subsite can be observed in Figure 4.

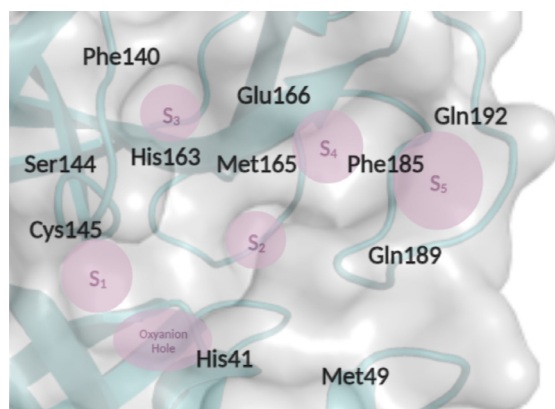


Figure 4. Catalytic site of SARS-CoV-2 M^{pro}. Each subsite is marked by a red circle and named, the key residues that form the subsite are shown together with the catalytic dyads, His41 and Cys145 [48]. Created with BioRender.com and PyMOL.

The specificity of interaction allows this protein to be able to recognize certain motifs on the peptide cleavage sites. The conservation of amino acid residues within the catalytic site in M^{pro} of SARS-CoV-2, SARS-CoV, and MERS-CoV allows these three enzymes to recognize the same motif, which is (L/F/M/P/V)Q – (S/A/G/N) [33, 56]. Understanding the surface geometry and chemical characteristics of the peptide substrate at a certain position that interacts with the corresponding subsite makes it possible to design a synthetic inhibitor compound. The inhibitor can be synthesized to have a struc-

ture with geometry and chemical properties that mimic the substrate peptide or commonly known as a peptidomimetic compound. Not only the peptidomimetic, small organic molecules that have surface characteristics that match the subsite on M^{pro} can also be considered potential inhibitors [57, 58]. Details of the interaction specificity between each subsite and its corresponding position within the substrate can be seen in Table 3.

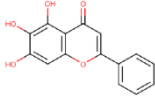
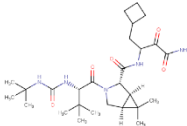
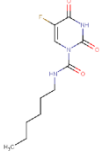
Table 3. Interaction specificity of each subsite [49].

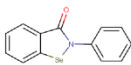
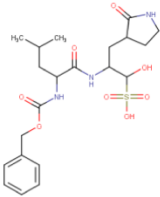
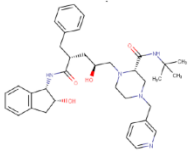
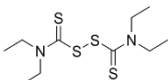
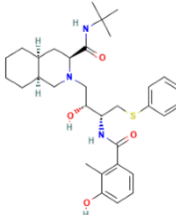
Subsite	Position	Key Residues	Substrate Specificity
S1	P1	Phe140, Gly143, Ser144, Cys145, His163, Glu166, His172	Gln, lactam-derivative
S2	P2	Thr25, His41, Cys145	Leu, Phe, Met, Val
S3	P3	His41, Met49, Met165	No specificity
S4	P4	Met165, Glu166	Hydrophobic acids
S5	P5	Glu166, Met165, Gln189	Hydrophobic acids

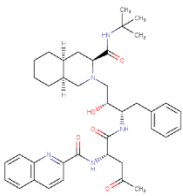
As an important part of SARS-CoV-2 M^{pro}, the catalytic site becomes the main target in the development of an anti-virus that works as an inhibitor for this enzyme. Inhibition of M^{pro} could prevent the maturation of new virion particles within the host cell, causing the interruption of SARS-CoV-2 replication cycle [59]. Many natural products and medicinal plants were studied as sources of inhibitory compounds and were later found to have excellent antiviral activity against the *Coronaviridae* family. Several classes of compounds such as Michael acceptor inhibitors, α -keto amides, aldehydes,

ketones, and their derivatives are known to possess the strongest interaction towards the catalytic site of M^{pro}. It means that these classes of compounds have the greatest potential to be developed as M^{pro} inhibitor [29]. Even the interaction with Michael acceptor inhibitors and epoxy peptidomimetic compounds can occur irreversibly because these sorts of compounds tend to be resistant to hydrolysis by forming covalent bonds with the thiol group of Cys145 which is located within the S1 of substrate-binding pocket [60]. Some examples of inhibitors targeting SARS-CoV-2 M^{pro} can be seen in Table 4.

Table 4. Example of inhibitors targeting M^{pro} SARS-CoV-2.

Drug Name & Structure	Compound Class	Original Function	%F	IC ₅₀	Ref.
Baicalein					
Drug Bank ID: DB16101 PubChem ID: 5281605	Flavonoid & polyflavonoid	Anti-cancer drug	3.61% $\pm$ 4.4%	5.158 $\pm$ 0.928 μ M	[79,80]
					
Boceprevir					
Drug Bank ID: DB08873 PubChem ID: 10324367	Peptidomimetic inhibitor	Hepatitis C virus protease inhibitor	In rats and dogs ranged between 24% to 34% Absolute %F is not known yet	4.1 $\pm$ 0.9 μ M	[81,82]
					
Carmofur					
Drug Bank ID: DB09010 PubChem ID: 2577	Pyrimidine derivative	Antineoplastic agent	95.5% $\pm$ 22.7%	1.82 $\pm$ 0.06 μ M	[83,84]
					

Ebselen Drug Bank ID: DB12610 PubChem ID: 3194						
	Benzenoid compound	Antioxidant	N.A.	30.91 ± 2.67 nM	[85]	
GC-376 Drug Bank ID: DB15796 PubChem ID: 13170475						
	Peptidomimetic inhibitor	Coronavirus M ^{pro} inhibitor	3%	0.03 to 0.16 μM	[86,87]	
Indinavir Drug Bank ID: DB00224 PubChem ID: 5362440						
	Peptidomimetic inhibitor	HIV-1 protease inhibitor	65%	43.1 ± 2.8 μM	[88,89]	
Disulfiram Drug Bank ID: DB00822 PubChem ID: 3117						
	Disulfide carbamate derivative	Alcohol antagonist	80% to 90%	4.7 ± 0.4 μM	[89,90]	
Nelfinavir Drug Bank ID: DB00220 PubChem ID: 64143						
	Peptidomimetic inhibitor	HIV-1 protease inhibitor	20% to 80%	234 μM	[91,93]	

Saquinavir Drug Bank ID: DB01232 PubChem ID: 441243					
					
Peptidomimetic inhibitor	HIV-1 protease inhibitor	Dose dependent	411 μ M	[92,93]	

SARS-CoV-2 M^{pro} is one of the potential target proteins for the development of antivirals for COVID-19 treatment besides the spike protein (S-protein). The SARS-CoV-2 M^{pro} even possesses a greater potential than S-protein, considering that S-protein is much more prone to mutations that will lead it to develop various variants [61, 62]. The rapid mutation rate of S-protein causes the efficacy of all drugs that have been developed to inhibit its activity to decrease over time [61]. Even though M^{pro} has a lower rate of mutation, a study that has been conducted by Ullrich, et al. in 2022 on some variants of SARS-CoV-2 showed that M^{pro} also undergoes mutations. These major mutations were found in C.37-Lambda, B.1.1.318, B.1.2, B.1.351-Beta, B.1.1.529-Omicron, and P.2-Zeta variant of SARS-CoV-2 and are considered as a missense mutation or non-synonymous mutation, which are G15S, T21I, L89F, P132H, and L205V respectively [61]. Enzyme kinetics studies showed that the variants of M^{pro} have a catalytic activity that does not differ much from the wild type. This suggests that the mutations found in M^{pro} do not occur in an essential part of the protein for supporting the enzymatic activity. The conservation of these parts can be used as the main target for SARS-CoV-2 M^{pro} inhibitor development [61].

Another study conducted by Krishnamoorthy et al. in 2021 revealed that there are 282 amino acid residues of M^{pro} SARS-CoV-2 that undergo nonsynonymous mutation. The study was done by analyzing the mutation rate of each residue from 19,154 samples of mutant M^{pro} sequence which sized exactly 306 amino acid residues. The position of amino acid residues that undergo mutation, even if it is only one sequence, is known as a mutational hotspot or hotspot residue [67, 68]. The other 24 amino acid residues were later known as mutational coldspots or coldspot residues, which are the

residues that had not undergone mutation yet at the time of data collection. There were found that 15 coldspot residues are located in domain I and domain II, while the other 9 were found in domain III. However, it is not a zero possibility that mutation can occur in these amino acid residues in the future [67, 69]. In addition, mutational coldspots show a great resistance of amino acids to mutations. It also can be said that these coldspot residues have a greater conservation rate. The development of SARS-CoV-2 M^{pro} inhibitors can be carried out by targeting the coldspot residues to avoid the possibility of the effectivity-loss of the inhibitors due to mutations [67].

However, the catalytic site is not free from mutation. It was found that a total of 25 amino acid residues make up the catalytic site and it was later known that 22 of these residues have been reported to undergo mutation, while the 3 amino acid residues are seen to be the coldspot residues, which are Leu141, Phe185, and Gln192. Some coldspot residues apart from these three can also be observed within domain I as well as domain II, and some of them are located near the substrate-binding pocket [67]. The coldspot Asn133 and Lys137 were found beneath the surface of the substrate-binding pocket. The coldspot residues are even involved in the substrate-binding site and interact with Phe185 and Gln192 to stabilize the catalytic site of M^{pro}. In addition, the coldspots Leu27, Asn119, and Gly146 are even located near the catalytic dyads of this enzyme. This suggests that these amino acid residues are involved in the core of the catalytic reaction. It is even known that Leu27 plays an important role in the catalytic activity of M^{pro} and together with Asn119 it helps to form the binding site [67].

4. Allosteric Site of SARS-CoV-2 M^{pro}

4.1 Distal Site

To date, it is only known that six allosteric sites can be targeted for the development of M^{pro} inhibitors. These allosteric sites were located on the surface of each protomer and were discovered together with eight different ligands that co-cryst-

tallized with the sites [34]. Analysis of the druggability score of each site was then carried out based on the pocket size, hydrophobicity, and accessibility of the site. It was later discovered that only two of these allosteric sites showed very high scores of druggability (DScore > 0.90), which were the allosteric 2 (DScore = 1.02) and the allosteric 5 (DScore =

1.11) [34, 68]. However, allosteric 3 was known not to give a good score while the other allosteric sites were not detected as druggable sites. This could be happened due to the size of the pocket within a site is relatively small and shallow, so that the ligand cannot bind stably to the sites [34, 71].

Besides targeting the catalytic site of M^{pro}, which is the largest ligand-binding pocket of this protein, the development of M^{pro} inhibitors can also be carried out by targeting the allosteric sites such as distal sites and dimerization interface [34, 72]. Computational studies have predicted that targeting both distal sites and dimerization interface will result in allosteric modulation of the M^{pro} protomer, including conformational changes or thermal fluctuations of the protomer [65, 73]. Analysis of key residues of M^{pro} that interact with ligands indicated the presence of a pocket at the interface between domain II and domain III. This pocket that is located quite far away from the catalytic site is later known as the distal sites [34].

The distal site is found in the area near the C-terminal region and is known to have high druggability. Several factors that are affecting this high score are this distal site has key residues to interact with ligands and this site could promote the allosteric modulation in the form of interdependent dynamic effects on the active site, which is located near the N-finger. In other words, the interaction between the distal site with the inhibitor could indirectly induce conformational changes to the active site which can reduce the activity of SARS-CoV-2 M^{pro} [65]. This distal site contains various amino acid residues that make up the β -sheet and α -helix. If there is an inhibitor that can interact thermodynamically stable, then it will be a reversible noncompetitive inhibitor. Currently, no inhibitor has been reported to work on the distal site. Evaluation that has been done by Alzyoud et al in 2022 using PockDrug revealed that the pocket's druggability score was in the range of 0.37 to 1.00 where 60% of which had a value above 0.90. This indicates the great potential of the distal site as a target for allosteric inhibitor of M^{pro} [34].

Nevertheless, mutations occurred on distal sites could affect the intramolecular stability of each SARS-CoV-2 M^{pro} protomer. For example, the mutant P132H which is found in the SARS-CoV-2 variant of B.1.1.529 Omicron. This mutation is located in turn between two β -sheet of M^{pro}, quite far away from the substrate-binding site, and causes a shift in the melting temperature point (T_m) of the enzyme [56, 58, 59]. This is due to a slight conformational alteration caused by His132. The imidazole group of His132 will undergo protonation when it is around physiological pH and makes this residue becomes positively charged [63, 65]. This can induce the Glu240 residue within domain III of M^{pro} to interact with His132 through electrostatic interactions. In addition, the water molecule between the two amino acids can mediate the formation of a hydrogen bond between His132 and Glu240. Thr198, which is located in the area near these two amino

acids, will undergo rotation for about 90° and its hydroxyl group will be in close proximity to Glu240 to support the hydrogen bond [63]. Overall, there will be an adjustment of the rotamer of the other amino acid residues within the protomer due to the mutation of P132H. This is what causes the decrease in the intramolecular stability of the protomer [63, 66]. However, it is also known that the P132H mutation does not cause a decrease in the enzymatic activity of M^{pro}. Instead, the mutation is only affecting the flexibility of the protomer by increasing its intramolecular kinetic energy. This can give a major impact on the evolution of M^{pro} by allowing this enzyme to recognize broader spectrum of substrate profiles as well as increasing the ability of this enzyme to fit the geometry of the substrate much better [63].

4.2 Dimerization Site

M^{pro} of SARS-CoV-2 is a type of protease that can only be active when it is in a homodimeric state. The interactions that are formed between two identical protomers in the homodimeric structure allow each protomer of M^{pro} to be active to catalyze the proteolysis of substrate peptide [54, 74]. However, due to the complexity of interactions within the homodimeric structure, the two protomers are rarely active simultaneously at the same time, instead, they will be active alternately. Nevertheless, both protomers will be inactive if they do not interact with each other because the substrate-binding pocket of one protomer requires interaction with the counterpart protomer to be opened and accessible by the peptide substrate [46, 75]. Even though the substrate peptide has managed to access the substrate-binding pocket when the protomer is in a single state, the interactions that occur tend to be weak and unstable. This causes the single protomer of SARS-CoV-2 M^{pro} to possess very low, even almost no proteolytic activity at all [46, 66, 75].

Dimerization involves the interaction between the amino acid residues of the N-finger of one protomer and the amino acid residues of the counterpart protomer. However, the residues of dimerization are scattered all around the dimerization interface of SARS-CoV-2 M^{pro} which involves all parts of this protein [46, 75]. Even some of the dimerization residues in domain II are the residues that also make up the S1 of the substrate-binding pocket, such as Phe140 and Glu166 [44, 73]. The interactions are known to assist the opening of substrate-binding pocket S1, so that it becomes accessible by the substrate peptides and could also increase the overall stability of the catalytic site [46, 75].

In addition to the interaction between the N-finger of one protomer and the dimerization residues around the substrate binding pocket of another protomer, there is another factor in the SARS-CoV-2 M^{pro} dimerization process that could help increase the stability and promote the activation of each protomer [64, 74]. This factor is related to the intramolecular stability of each protomer when in a monomeric state and

homodimeric state. It is known that the protomer of SARS-CoV-2 M^{pro} has very low intramolecular stability when in its monomeric or single state. It means that the overall molecule of protomer could undergo significant conformational changes within a short period of time [66, 76].

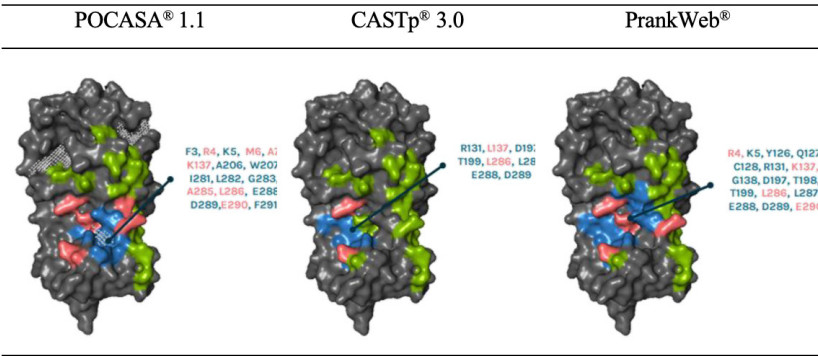
This rapid change of conformation is caused by the high kinetic energy of the atoms within the protomer molecule. This condition will decrease the affinity of the protomer's catalytic site to the substrate peptide which will lead to the inactivity of the protomer. Moreover, the proteolytic reaction cannot occur if the substrate peptide is not bound properly and stably to the substrate-binding pocket of the protomer due to rapid changes in the geometry of the substrate-binding pocket [66, 76, 77]. Interaction between these two protomers in order to form the homodimeric structure could also increase the intramolecular stability of each protomer. This is because the interaction can help to reduce the kinetic energy of the atoms within each protomer [66, 76, 77]. The increase in intramolecular stability allows the protomer to interact with the peptide substrate much more stable and the proteolytic reaction can occur easily. This dimerization is later stabilized by the interaction of each domain III from both protomers [66, 76, 77].

The dimerization of the two SARS-CoV-2 M^{pro} protomers involves 26 amino acid residues which are scattered over almost all parts of the SARS-CoV-2 M^{pro}. These residues tend to be resistant to mutation which means that they are relatively conserved [50]. The further study provides information that the dimerization interface of each protomer can be divided into three regions of dimerization. The first region is located at the N-terminus of the protomer, spanning the entire N-finger and a small portion of the domain I which extends from Ser1 to Cys16. The interactions that occur in this region are including Ser1 with Phe140' and Glu166' (from the counterpart protomer), Gly2 with Phe140', Arg4 with Lys137' and Glu290', Met6 with Glu290', Ala7 with Val125', Phe8 with

Val125', Ser10 with Ser10', Gly11 with Gly11', and Glu14 with Glu14' [53]. The second dimerization region is located on the β -strand of domain II and extends from Tyr118 to Val125. The interactions that occur in this region are including Asn119 with Phe305' which is mediated by water molecules, Pro122 with Phe305', and Val125 with Ala7' [53]. The third dimerization region is located in a loop that extends from Lys137 to Ser144. The interactions that occur in this region are including Lys137 with Arg4', Ser139 with Gln299', Phe140 with Gly2', and Leu141 with Ser1' [53]. The interactions between Ala285 with Ala285', Leu286 with Leu286', Asp295 with Asp295', and Gln299 with Gln299' in domain III must also be considered because the interaction of these residues helps to increase the stability of the homodimeric structure of SARS-CoV-2 M^{pro} [45]. Based on this information, the dimerization residues are Ser1, Gly2, Arg4, Met6, Ala7, Phe8, Ser10, Gly11, Glu14, Asn28, Asn119, Pro122, Ser123, Val125, Lys137, Ser139, Phe140, Ser147, Glu166, Ala285, Leu286, Glu290, Asp295, Arg298, Gln299, and Phe305 which total of amino acid residues involved in the formation of the homodimer is 26 residues [45, 50, 53].

The apo-structure of the SARS-CoV-2 M^{pro} protomer does not have a clear ligand-binding pocket on its dimerization interface because this part is involved in increasing the dimerization stability of the protein as well as the active site. However, when in a monomeric state (single protomer), M^{pro} can show a good ligand-binding pocket with a size and polarity that allows it to bind a ligand [34, 71]. This indicates that the monomer of SARS-CoV-2 M^{pro} tends to have better druggable pockets than the apo-structure one [34, 71]. The ligand-binding pocket for the dimerization interface of each SARS-CoV-2 M^{pro} protomer has already been predicted by using three different softwares by Azmi in 2021 (nonpublished BSc. research thesis) [78] as can be seen in Table 5.

Table 5. Predicted ligand-binding pocket in SARS-CoV-2 M^{pro} dimerization interface. Color details: Green = Dimerization residues, Blue = Dimerization residues that contribute to the formation for the ligand binding pocket, Pink = Non-dimerization residues that contribute to the formation of the ligand-binding pocket, Grey = Remaining residues in SARS-CoV-2 M^{pro}. Protein structure shown was the 3D model of 6LU7 from PDB [78].



In the context of mutations, half of the known coldspots are located on the surface of M^{pro} while the rest are buried within the 3D structure of the protein. It is observed that there are 7 coldspot residues on the surface of M^{pro} which are involved in the dimerization process, which are Gly2, Arg4, Tyr126, Lys137, Leu141, Leu286, and Leu287. These coldspots form two dimerization sites [67]. The first site involves Gly2, Arg4, Tyr126, Lys137 and Leu141. The second site involves Arg4, Lys137, Leu286, and Leu287. Even the interaction at these two sites is a key interaction that plays an important role in maintaining the stability of the homodimer structure and the active site [67]. Hydrogen bond occurs between Arg4 and Lys137 with three other coldspots around it (Gly2, Tyr126, and Leu141). This makes these amino acids become a very potential target in the development of dimerization inhibitors of SARS-CoV-2 M^{pro} [67].

5. Concluding Remarks

It is possible to create an antiviral by targeting the M^{pro} catalytic or allosteric sites. The catalytic site, which is the protein's primary substrate-binding pocket, is where the proteolysis reaction of substrate peptide takes place. The function of M^{pro} is lost when there are antivirals that target the catalytic region and act as competitive inhibitors against substrate peptides, preventing each protomer from interacting with the substrate peptide. In addition to the catalytic site, the surface of M^{pro} protomer also has some ligand-binding pockets that can bind substrates, which are known as allosteric sites. In general, this structure is further divided into the distal site and the dimerization interface. Antiviral that targets the distal site will alter the 3D structure of the protein, resulting in allosteric modulation of the protomer. The shape and surface structure of the catalytic site will be indirectly affected by this, which will reduce the affinity of the substrate peptide for the M^{pro} catalytic site. Antivirals that function at the dimerization interface will, however, prevent the two protomers of M^{pro} from forming an active homodimer structure. The protomer's proteolytic activity is so minute as to be regarded as negligible when the protomer is in a single state. M^{pro} SARS-CoV-2 will become completely inactive as a result. Mutational studies on the structure of SARS-CoV-2 M^{pro} also show that this protein tends to be resistant to mutations. It was also discovered that this protein had 24 mutational coldspot residues, which may serve as the major target for the development of M^{pro} inhibitors. The incidence of mutations within amino acid residues that are known to undergo mutations was likewise found to be extremely low, and the changes that did take place had no impact on the protein's ability to function. This suggests that SARS-CoV-2 M^{pro} has great potential to be the candidate for SARS-CoV-2 antiviral drug development.

Acknowledgement

We acknowledge support of this work by the project "Dimer-Based Screening System (DBSS) Implementation for Anti-COVID-19 Candidates Screening From Indonesian Biodiversity", which is implemented under the Program "Research and Innovation for Developed Indonesia", funded by the "Indonesia Endowment Fund for Education", according to the Decree of the Deputy for National Research and Innovation Facilitation, National Research and Innovation Institute, No. 60/II/HK/2022. The authors declare no conflict of interest.

References

- [1.] Damme, W., Dahake, R., Delamou, A., et al. (2020). The COVID-19 pandemic: diverse context; different epidemics – how and why. *BMJ Journal for Global Health Analysis*, 5: 1-16. DOI: <https://doi.org/10.1136/bmjgh-2020-003098>
- [2.] Balkhair, A. A. (2020). COVID-19 pandemic: a new chapter in the history of infectious disease. *Oman Medical Journal*, 35(2): 44-45. DOI: <https://doi.org/10.5001/omj.2020.41>
- [3.] Helmy, Y. A., Fawzy, M., Elasad, A., et al. (2020). The COVID-19 pandemic: a comprehensive review of taxonomy, genetics, epidemiology, diagnosis, treatment, and control. *MDPI Journal of Clinical Medicine*, 9(1225): 2-29. DOI: <https://doi.org/10.3390/jcm9041225>
- [4.] Worldometer. (2021). COVID-19 Coronavirus Pandemic, Last Updated September 23, 2022, 08.31 GMT [online] accessed from <https://www.worldometers.info/coronavirus/> on September 23, 2023 at 23.30 CET
- [5.] Ritchie, H., Ospina, E. O., Beltekian, D.; et al. (2021). Coronavirus (COVID-19) cases. *Our World in Data* [online] accessed from <https://ourworldindata.org/covid-cases> on September 23, 2022 at 21.30 CET
- [6.] Raphael, T. & Fazeli, S.. (2022). COVID's Fifth Wave Shows Us How to Live With the Virus [online] accessed from <https://www.bloomberg.com/opinion/articles/2022-03-22/covid-s-fifth-wave-shows-us-how-to-live-with-the-omicron-virus-subvariant> on September 23, 2023 at 21.30 CET
- [7.] Rehman, M. F., Fariha, C., Anwar, A., et al. (2021). Novel coronavirus disease (COVID-19) pandemic: a recent mini review. *Computational and Structural Biotechnology Journal*, 19: 612-623. DOI: <https://doi.org/10.1016/j.csbj.2020.12.033>
- [8.] Li, Q., Wu, J., Nie, J., et al. (2020). The impact of mutations in SARSCoV-2 spike on viral infectivity and antigenicity. *Elsevier Journal for Cell*, 182:1284-1294. DOI: <https://doi.org/10.1016/j.cell.2020.07.012>
- [9.] Gonzalez, M. L., Ortiz, M. S., & Landete, P. (2022). Evolution and clinical trend of SARS-CoV-2 variants. *Open Respiratory Archives*, 4: 1-3. DOI: <https://doi.org/10.5614/3bio.2023.5.2.2>

- org/10.1016/j.opresp.2022.100169
- [10.] Gorbalenya, A. E., Baker, S. C., Baric, R. S., et al. (2020). The species Severe Acute Respiratory Syndrome – Related Coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. *Nature Journal of Microbiology*, 5(4): 536-544. DOI: <https://doi.org/10.1038/s41564-020-0695-z>
 - [11.] Cevik, M., Kuppalli, K., Kindrachuk, J., Peiris, M.. (2020). Virology, transmission, and pathogenesis of SARS-CoV-2. *BMJ Journal of Virology*, 371: 1-6. DOI: <https://doi.org/10.1136/bmj.m3862>
 - [12.] Hartenian, E., Nandakumar, D., Lari, A., et al. (2020). The molecular virology of coronaviruses. *Journal of Biology and Chemistry*, 295(37): 12910-12934. DOI: <https://doi.org/10.1074/jbc.REV120.013930>
 - [13.] Shah, K.R., Utekar, D.B., Nikam, S.S., et al. (2020). The emergent pandemic – a review on coronavirus SARS-CoV-2: virology, pathogenesis and outbreak. *Journal of Infectious Diseases and Epidemiology*, 6(3): 1-8. DOI: <https://doi.org/10.23937/2474-3658/1510136>
 - [14.] Gennaro, F.D., Pizzol, D., Marotta, C., et al. (2020). Coronavirus Diseases (COVID-19) current status and future perspectives: a narrative review. *MDPI International Journal of Environmental Research and Public Health*, 17(2690): 1-11. DOI: <https://doi.org/10.3390/ijerph17082690>
 - [15.] Park, Su Eun. (2020). Epidemiology, virology, and clinical features of severe acute respiratory syndrome-coronavirus-2 (SARS-CoV-2; Coronavirus Disease-19). *CEP Journal of Virology*, 63(4): 119-124. DOI: <https://doi.org/10.3345/cep.2020.00493>
 - [16.] Zheng, Jun. (2020). SARS-CoV-2: an emerging coronavirus that causes a global threat. *International Journal of Biological Sciences*, 16(10): 1678-1685. DOI: <https://doi.org/10.7150/ijbs.45053>
 - [17.] Poltronieri, P., Sun, B., & Mallardo, Massimo. (2015). RNA viruses: RNA roles in pathogenesis coreplication and viral load. *Current Genomics*, 16(5): 327-335. DOI: <https://doi.org/10.2174/1389202916666150707160613>
 - [18.] Velthuis, A. J. W.. (2014). Common and unique features of viral RNA-dependent polymerases. *Cellular and Molecular Life Sciences*. 71: 4403-4420. DOI: <https://doi.org/10.1007/s00018-014-1695-z>
 - [19.] Cirrincione, L., Plescia, F., Ledda, C., et al. (2020). COVID-19 pandemic: prevention and protection measures to be adopted at the workplace. *MDPI Journal for Sustainability*, 12(3603): 1-18. DOI: <https://doi.org/10.3390/su12093603>
 - [20.] Combe, M. & Sanjuan, R. (2014). Variation in RNA Virus Mutation Rates across Host Cells. *PLOS Pathogens*, 10(1): 1-7. DOI: <https://doi.org/10.1371/journal.ppat.1003855>
 - [21.] Monttinen, H. A. M., Ravantti, J. J., & Poranen, M. M. (2021). Structure unveils relationships between RNA virus polymerases. *MDPI Journal of Viruses*, 13(313): 1-17. DOI: <https://doi.org/10.3390/v13020313>
 - [22.] Sonntag, K. C., Darai, G. (1995). Evolution of viral DNA-dependent RNA polymerases. *Virus Genes*, 11: 271-284. DOI: <https://doi.org/10.1007/BF01728665>
 - [23.] Hillen, H. S., Kokic, G., Farnung, L., et al. (2020). Structure of replicating SARS-CoV-2 polymerase. *Nature*, 584: 154-156. DOI: <https://doi.org/10.1038/s41586-020-2368-8>
 - [24.] Le, T.T., Andreadakis, Z., Kumar, A., et al. (2020). The COVID-19 vaccine development landscape. *Nature Reviews for Drug Discovery*, 19: 305- 306. DOI: <https://doi.org/10.1038/d41573-020-00073-5>
 - [25.] Motyan, J., Mahdi, M., Hoffka, G., & Tozser, J.. (2022). Potential Resistance of SARS-CoV-2 Main Protease (Mpro) against Protease Inhibitors: Lessons Learned from HIV-1 Protease. *MDPI International Journal of Molecular Sciences*, 23(7): 3507. DOI: <https://doi.org/10.3390/ijms23073507>
 - [26.] Guo, S., Liu, K., & Zheng, J.. (2021). The genetic variant of SARS-CoV-2: would it matter for controlling the devastating pandemic?. *International Journal of Biological Sciences*, 17(6): 1476-1485. DOI: <https://doi.org/10.7150/ijbs.59137>
 - [27.] Forni, G., Mantovani, A., et al. (2021). COVID-19 vaccines: where we stand and challenges ahead. *Springer Nature Journal for Cell Death & Differentiation*, 28: 626-639. DOI: <https://doi.org/10.1038/s41418-020-00720-9>
 - [28.] Elshabrawy, Hatem A. (2020). SARS-CoV-2 : an update on potential antivirals in light of SARS-CoV antiviral drug discoveries. *MDPI Vaccines*, 8(2): 1-33. DOI: <https://doi.org/10.3390/vaccines8020335>
 - [29.] Palanisamy, K., Rubavathy, S. M. E., Prakash, M., et al. (2022). Antiviral activities of natural compounds and ionic liquids to inhibit the Mpro of SARS-CoV-2: A computational approach. *RSC Advances*, 6(2022). DOI: <https://doi.org/10.1039/D1RA08604A>
 - [30.] Kokic, G., Hillen, H. S., Tegunov, D., et al. (2021). Mechanism of SARS-CoV-2 polymerase stalling by remdesivir. *Nature Communications*, 12(279): 1-7. DOI: <https://doi.org/10.1038/s41467-020-20542-0>
 - [31.] Ferdiansyah, A., Nainu, F., Kuldeep, D., et al. (2021). Ramdesivir and its antiviral activity against COVID-19: A systematic review. *Clinical Epidemiology and Global Health*, 9: 123-127. DOI: <https://doi.org/10.1016/j.cegh.2020.07.011>
 - [32.] Gordon, D.E., Jang, G.M., Bouhaddou, M., et al. (2020). A SARS-CoV2 protein interaction map reveals targets for drug repurposing. *Nature*, 583: 459- 468. DOI: <https://doi.org/10.1038/s41586-020-2286-9>
 - [33.] Ullrich, Sven & Nitsche, Christoph. (2020). The SARS-CoV-2 main protease as drug target. *Elsevier Journal*

- for Bioorganic & Medicinal Chemistry Letters, 30: 1-8. DOI: <https://doi.org/10.1016/j.bmcl.2020.127377>
- [34.] Alzyoud, L., Ghattas, M. A., Atatreh, N.. (2022). Allosteric binding sites of the SARS-CoV-2 main protease: Potential targets for broad-spectrum anti-coronavirus agents. *Drug Design, Development and Therapy*, 16: 2463-2478. DOI: <https://doi.org/10.2147/DDDT.S370574>
- [35.] Flynn, J.M., Samant, N., Schneider-Nachum, G., Barkan, D.T., Yilmaz, N.K., Schiffer, C.A., Moquin, S.A., Dova-la, D., Bolon, D.N.A. (2022). Comprehensive fitness landscape of SARS-CoV-2 Mpro reveals insights into viral resistance mechanisms. *eLife*, 11, 1-27. DOI: <https://doi.org/10.7554/eLife.77433>
- [36.] Jin, Z. Du, X., Xu, Y., et al. (2020). Structure of Mpro from SARS-CoV-2 and discovery of its inhibitors. *Nature*, 582, 289-293. DOI: <https://doi.org/10.1038/s41586-020-2223-y>
- [37.] Rungruangmaitree, R., Phoochaijaroen, S., Chimprasit, A., Saparpakorn, P., Pootanakit, K., Tanramluk, D. (2023). Structural analysis of coronavirus main protease for the design of pan-variant inhibitors. *Sci Rep*, 13(7055). DOI: <https://doi.org/10.1038/s41598-023-34305-6>
- [38.] Firouzi, R., Ashouri, M., Karimi-Jafari, M.H. (2021). Structural insights into the substrate-binding site of main protease for the structure-based COVID-19 drug discovery. *Proteins: Structure, Function, and Bioinformatics*, 90(5), 1090-1101. DOI: <https://doi.org/10.1002/prot.26318>
- [39.] Henghasatporn, K., Harada, R., Wilasluck, P., et al. (2022). Promising SARS-CoV-2 main protease inhibitor ligand-binding modes evaluated using LB-PaCS-MD/FMO. *Sci Rep* 12, 17984. DOI: <https://doi.org/10.1038/s41598-022-22703-1>
- [40.] Chan, H.T.H., Moesser, M.A., Walters, R., et al. (2021). Discovery of SARS-CoV-2 Mpro peptide inhibitors from modelling substrate and ligand binding. *Chemical Science*, 12(13686). DOI: <https://doi.org/10.1039/d1sc03628a>
- [41.] Stoddard, S.V., Stoddard, S.D., Oelkers, B.K. (2020). Optimization rules for SARS-CoV-2 Mpro antivirals: Ensemble docking and exploration of the coronavirus protease active site. *Viruses*, 12(9), 942. DOI: <https://doi.org/10.3390/v12090942>
- [42.] Mohamadian, M., Chiti, H., Shoghli, A., et al. (2020). COVID-19: virology, biology and novel laboratory diagnosis. *Journal of Metabolic Disease*: 1-29. DOI: <https://doi.org/10.1002/jgm.3303>
- [43.] Gordon, D.E., Jang, G.M., Bouhaddou, M., et al. (2020). A SARS-CoV2 protein interaction map reveals targets for drug repurposing. *Nature*, 583: 459- 468. DOI: <https://doi.org/10.1038/s41586-020-2286-9>
- [44.] Estrada, E.. (2020). Topological analysis of SARS-CoV-2 main protease. *AIP Advances Journals of Mathematical Physics Collection*, 30: 1-14. DOI: <https://doi.org/10.1063/5.0013029>
- [45.] Zhang, L., Lin, D., Sun, X., et al. (2020). Crystal structure of SARSCoV-2 main protease provides a basis for design of improved alpha-ketoamide inhibitors. *Science Report Journal*: 1-8. DOI: <https://doi.org/10.1126/science.abb3405>
- [46.] Zhenming, J., Du, X., Xu, Y., et al. (2020). Structure of Mpro from SARSCoV-2 and discovery of its inhibitors. *Nature*, 582: 289-312. DOI: <https://doi.org/10.1038/s41586-020-2223-y>
- [47.] Firouzi, R., Ashouri, M., Jafari, M. H. K.. (2022). Structural insights into the substrate-binding site of main protease for the structure-based COVID-19 drug discovery. *Proteins: Structure, Function, and Bioinformatics*, 90(5): 1090-1101. DOI: <https://doi.org/10.1002/prot.26318>
- [48.] Sun, Z., Wang, L., Li, X., et al. (2022). An extended conformation of SARS-CoV-2 main protease reveals allosteric targets. *PNAS Journal of Biochemistry*, 119(15): 1-9. DOI: <https://doi.org/10.1073/pnas.2120913119>
- [49.] Gunther, S., Reinke, P. Y. A., Garcia, Y. F., et al. (2021). X-ray screening identifies active site and allosteric inhibitors of SARS-CoV-2 protease. *Science*, 372(6542): 642-646. DOI: <https://doi.org/10.1126/science.abf7945>
- [50.] Goyal, B., Goyal, D.. (2020). Targeting the dimerization of the main protease of coronaviruses: a potential broad-spectrum therapeutic strategy. *ACS Combinatorial Science*, 22(6): 297-305. DOI: <https://doi.org/10.1021/acscombsci.0c00058>
- [51.] Mensah, J. O., Ampomah, G. B., Gasu, E. N., et al. (2022). Allosteric modulation of the main protease (Mpro) of SARS-CoV-2 by Casticin – Insights from molecular dynamics simulations. *Chemistry Africa*, 488. DOI: <https://doi.org/10.1007/s42250-022-00411-7>
- [52.] Suarez, D., & Diaz, N. (2020). SARS-CoV-2 main protease: a molecular dynamics study. *Journal of Chemical Information and Modelling*, 60(12): 5815-5831. DOI: <https://doi.org/10.1021/acs.jcim.0c00575>
- [53.] Kneller, D. W., Phillips, G., O'Neill, H. M., et al. (2020). Structural plasticity of SARS-CoV-2 3CL Mpro active site cavity revealed by room temperature X-ray crystallography. *Nature Communications*, 11(3202): 1-6. DOI: <https://doi.org/10.1038/s41467-020-16954-7>
- [54.] Kneller, D. W., Phillips, G., Weiss, K. L.; et al. (2020). Unusual zwitterionic catalytic site of SARS-CoV-2 main protease revealed by neutron crystallography. *Journal of Biology and Chemistry*, 295(50): 17365-17373. DOI: <https://doi.org/10.1074/jbc.AC120.016154>
- [55.] Hu, Q., Xiong, Y., Zhu, G., et al. (2022). The SARS-CoV-2 main protease (Mpro): Structure, function, and emerging therapies for COVID-19. *Medical Communi-*

- cation: 1-27. DOI: <https://doi.org/10.1002/mco2.151>
- [56.] Monica, G. L., Bono, A., Lauria, A., et al. (2022). Targeting SARS-CoV-2 main protease for treatment of COVID-19: Covalent inhibitors structure – activity relationship insights and evolution perspectives. *Journal of Medicinal Chemistry*, 65(19): 12500-12534. DOI: <https://doi.org/10.1021/acs.jmedchem.2c01005>
- [57.] Pathak, M. K., Jha, V., Jain, N. K., et al. (2015). Review on peptidomimetics: A drug designing tool. *Indo American Journal of Pharmaceutical Research*, 5(12): 3859-3866
- [58.] Ko, E., Liu, J., Perez, L. M., et al. (2010). Universal peptidomimetics. *Journal of The American Chemical Society*, 133(3): 462-477. DOI: <https://doi.org/10.1021/ja1071916>
- [59.] Mahgoub, R., Mohamed F. E., Alzyoud, L., et al. (2022). The discovery of small allosteric and active site inhibitors of the SARS-CoV-2 main protease via structure-based virtual screening biological evaluation. *MDPI Journal of Molecules*, 27(6710): 1-21. DOI: <https://doi.org/10.3390/molecules27196710>
- [60.] Kunakbaeva, Z., Carrasco, R., & Rozas, I. (2003). An approximation to the mechanism of inhibition of cysteine proteases: nucleophilic sulphur addition to Michael acceptors type compounds. *Journal of Molecular Structure*, 626: 209-216. DOI: [https://doi.org/10.1016/S0166-1280\(03\)00086-1](https://doi.org/10.1016/S0166-1280(03)00086-1)
- [61.] Ullrich, S., Ekanayake, K. B., Otting, G., & Nitsche, C.. (2022). Main protease mutants of SARS-CoV-2 variants remain susceptible to nirmatrelvir. *Elsevier Public Health Emergency Collection : Bioorganic & Medicinal Chemistry Letters*, 62. DOI: <https://doi.org/10.1016/j.bmcl.2022.128629>
- [62.] Ng, T. I., Correia, I., Seagal, J., et al. (2022). Antiviral drug discovery for the treatment of COVID-19 infections. *MDPI Journal of Viruses*, 14(5): 1-27. DOI: <https://doi.org/10.3390/v14050961>
- [63.] Sacco, M. D., Hu, Y., et al. (2022). The P132H mutation in the main protease of Omicron SARS-CoV-2 decreases thermal stability without compromising catalysis or small molecule drug inhibition. *Nature of Cell Research*, 32: 498-500. <https://doi.org/10.1038/s41422-022-00640-y>
- [64.] Sedova, M., Jaroszewski, L., Iyer, M., et al. (2022). Monitoring for SARS-CoV-2 drug resistance mutations in broad viral populations. Pre-print. DOI: <https://doi.org/10.1101/2022.05.27.493798>
- [65.] Malinska, M., Dauter, M., Kowiel, M., et al. (2015). Protonation and geometry of histidine rings. *Acta Crystallographica Section D - Biological Crystallography*, 71(7): 1444-1454. <https://doi.org/10.1107/S1399004715007816>
- [66.] Greasley, S. E., Noell, S., Plotnikova, O., et al. (2022). Structural basis for the in vitro efficacy of nirmatrelvir against SARS-CoV-2 variants. *Journal of Biological Chemistry*, 298(6). DOI: <https://doi.org/10.1016/j.jbc.2022.101972>
- [67.] Krishnamoorthy, N. & Fakhro, K.. (2021). Identification of mutation resistance cold spots for targeting the SARS-CoV-2 main protease. *JUBMB Life*, 73:670-675. DOI: <https://doi.org/10.1002/iub.2465>
- [68.] Stromich, L., Wu, N., Barahona, M., et al. (2022). Allosteric hotspots in the main protease of SARS-CoV-2. *Journal of Molecular Biology*, 434: 1-15. DOI: <https://doi.org/10.1016/j.jmb.2022.167748>
- [69.] Padhi, A. K., & Tripathi, T. (2022). Hotspot residues and resistance mutations in the nirmatrelvir-binding site of SARS-CoV-2 main protease: design, identification, and correlation with global circulating viral genomes. *Biochemical and Biophysical Research Communications*, 629: 54-60. DOI: <https://doi.org/10.1016/j.bbrc.2022.09.010>
- [70.] Schmidtke, P., & Barril, X. (2010). Understanding and predicting druggability, a high-throughput method for detection of drug binding sites. *Journal of Medicinal Chemistry*, 53(15): 5858-5867. DOI: <https://doi.org/10.1021/jm100574m>
- [71.] Alzyoud, L., Bryze, R. A., Al-Sorkhy, M., et al. (2022). Structure-based assessment and druggability classification of protein-protein interaction sites. *Nature Scientific Reports*, 12(7975): 1-18. DOI: <https://doi.org/10.1038/s41598-022-12105-8>
- [72.] Kubra, B., Badshah, S. L., Faisal, S., et al. (2022). Inhibition of the predicted allosteric site of the SARS-CoV-2 main protease through flavonoids. *Journal of Biomolecular Structure and Dynamics*. 41(18): 9103-9120. DOI: <https://doi.org/10.1080/07391102.2022.2140201>
- [73.] Samrat, S. K., Xu, J., Xie, X., et al. (2022). Allosteric inhibitors of the main protease of SARS-CoV-2. *Elsevier Public Health Emergency Collection*, 205(105381). DOI: <https://doi.org/10.1016/j.antiviral.2022.105381>
- [74.] Ferreira, J. C., Fadl, S., & Rabeh, W. M. (2022). Key dimer interface residues impact the catalytic activity of 3CLpro, the main protease of SARS-CoV-2. *Journal of Biological Chemistry*, 298(6). DOI: <https://doi.org/10.1016/j.jbc.2022.102023>
- [75.] Zhenming, J., Zhao, Y., Sun, Y., et al. (2020). Structural basis for the inhibition of SARS-CoV-2 main protease by antineoplastic drug carmofur. *Nature Structural & Molecular Biology*. 27: 529-532. DOI: <https://doi.org/10.1038/s41594-020-0440-6>
- [76.] Silvestrini, L., Belhaj, N., Comez, L., et al. (2021). The dimer-monomer equilibrium of SARS-CoV-2 main protease is affected by small molecule inhibitors. *Nature Scientific Reports*, 11(9283): 1-16. DOI: <https://doi.org/10.1038/s41598-021-88630-9>
- [77.] Reynolds, C. H. & Halloway, M., K. (2011). Thermodynamics of ligand binding and efficiency. *ACS Medic-*

- inal Chemistry Letters, 2(6): 433-437. DOI: <https://doi.org/10.1021/ml200010k>
- [78.] Azmi, M.H.S. (2021). In silico Drug Repurposing Studies In The Design of Cross-Linkin Assay and Dimer Based Screening System as A Screening System for SARS-CoV-2 Main Protease Inhibitor Development. Translated from the original title in Bahasa Indonesia: “Studi Drug Repurposing Secara In Silico dalam Perancangan Sistem Penapisan Kandidat Inhibitor Dimerisasi Main Protease SARS-CoV-2 untuk Cross-Linking Assay dan Dimer-Based Screening System”.
- [79.] Xiao, T., Cui, M., Zheng, C., Zhang, P., Ren, S., Bao, J., Gao, D., Sun, R., Wang, M., Lin, J., Zhang, L., Li, M., Li, D., Zhou, H., Yang, C. (2021). Both baicalein and gallic acid effectively inhibit SARS-CoV-2 replication by targeting Mpro and sepsis in mice. *Inflammation* 45, 1076-1088. DOI: <https://doi.org/10.1007/s10753-021-01602-z>
- [80.] Zhu, B., Zhang, Q., Wang, J.R., Mei, X. (2017). Cocystals of baicalein with higher solubility and enhanced bioavailability. *Crystal Growth & Design* 17(4), 1893-1901. DOI: <https://doi.org/10.1021/acs.cgd.6b01863>
- [81.] Howe, A.Y.M. & Ventkatraman, S. (2013). The discovery and development of boceprevir: A novel, first-generation inhibitor of the hepatitis C virus NS3/4A serine protease. *Journal of Clinical and Translational Hepatology*, 1(1), 22-32. DOI: <https://doi.org/10.14218/JCTH.2013.002XX>
- [82.] Göhl, M., Zhang, L., El-Kilani, H., Sun, X., Zhang, K., Brönstrup, M., Hilgenfeld, R. (2022). From repurposing to redesign: optimization of boceprevir to highly potent inhibitors of the SARS-CoV-2 Main Protease. *Molecules*, 27(13), 4292. DOI: <https://doi.org/10.3390/molecules27134292>
- [83.] Zhuang Z., Zhu, H., Wang, J., Zhu, M., Wang, H., Pu, W., Bian, H., Chen, L., Zhang H. (2013). Pharmacokinetic evaluation of novel oral fluorouracil antitumor drug S-1 in Chinese cancer patients. *Acta Pharmacologica Sinica*, 34(4), 570-580. DOI: <https://doi.org/10.1038/aps.2012.169>
- [84.] Jin, Z., Zhao, Y., Sun, Y., et al. (2020). Structural basis for the inhibition of SARS-CoV-2 main protease by anti-neoplastic drug carmofur. *Nature Structural & Molecular Biology*, 27, 529-532. DOI: <https://doi.org/10.1038/s41594-020-0440-6>
- [85.] Zmudzinski, M., Rut, W., Olech, K., et al. (2023). Ebselen derivatives inhibit SARS-CoV-2 replication by inhibition of its essential proteins: PLpro and Mpro proteases, and nsp14 guanine N7-methyltransferase. *Scientific Report*, 13, 9161. DOI: <https://doi.org/10.1038/s41598-023-35907-w>
- [86.] Joyce, R.P., Hu, V.W., Wang, J. (2022). The history, mechanism, and perspectives of nirmatrelvir (PF-07321332): and orally bioavailable main protease inhibitor used in combination with ritonavir to reduce COVID-19 related hospitalization. *Medicinal Chemistry Research*, 31, 1637-1646. DOI: <https://doi.org/10.1007/s00044-022-02951-6>
- [87.] Klacsová, M., Čelková, A., Búcsi, A., et al. (2022). Interaction of GC376, a SARS-CoV-2 Mpro inhibitor, with model lipid membranes. *Colloids Surf B Biointerfaces*, 220, 112918. DOI: <https://doi.org/10.1016/j.colsurfb.2022.112918>
- [88.] Kappelhoff, B.S., Huitema, A.D.R., Sankasting, S.U.C., et al. (2005). Population pharmacokinetics of indinavir alone and in combination with ritonavir in HIV-1 infected patients. *British Journal of Clinical Pharmacology*, 60(3), 276-286. DOI: <https://doi.org/10.1111/j.1365-2125.2005.02436.x>
- [89.] Li, Z., Li, X., Huang, Y.Y., et al. (2020). Identify potent SARS-CoV-2 main protease inhibitors via accelerated free energy perturbation-based virtual screening of existing drugs. *Biophysics and Computational Biology*, 117(44), 27381-27387. DOI: <https://doi.org/10.1073/pnas.2010470117>
- [90.] Lanz, J., Biniarz-Harris, N., Kuvaldina, M., et al. (2023). Disulfiram: Mechanism, applications, and Challenges. *Antibiotics*, 12(3), 524. DOI: <https://doi.org/10.3390/antibiotics12030524>
- [91.] Pai, V.B., & Nahata, M.C. (1999). Nelfinavir mesylate: a protease inhibitor. *Annals of Pharmacotherapy*, 33(3). DOI: <https://doi.org/10.1345/aph.18089>
- [92.] Facklam, M.M., Burhenne, J., Ding, R., et al. (2002). Dose-dependent increase of saquinavir bioavailability by the pharmaceutic aid cremophor EL. *British Journal of Clinical Pharmacology*, 53(6), 576-581. DOI: <https://doi.org/10.1046/j.1365-2125.2002.01595.x>
- [93.] Antonopoulou, I., Sapountzaki, E., Rova, U., Christakopoulos, P. (2022). Inhibition of the main protease of SARS-CoV-2 (Mpro) by repurposing/designing drug-like substances and utilizing nature's toolbox of bioactive compounds. *Computational Structural Biotechnology Journal*, 20, 1306-1344. DOI: <https://doi.org/10.1016/j.csbj.2022.03.009>