

RESEARCH

Open Access



Crystal structures of main protease (M^{Pro}) mutants of SARS-CoV-2 variants bound to PF-07304814

Haihai Jiang^{1†}, Xiaofang Zou^{2†}, Pei Zeng^{3,4}, Xiangyi Zeng^{3,4}, Xuelan Zhou^{3,4}, Jie Wang^{3,4}, Jin Zhang^{1*} and Jian Li^{2*}

Abstract

There is an urgent need to develop effective antiviral drugs to prevent the viral infection caused by constantly circulating SARS-CoV-2 as well as its variants. The main protease (M^{Pro}) of SARS-CoV-2 is a salient enzyme that plays a vital role in viral replication and serves as a fascinating therapeutic target. PF-07304814 is a covalent inhibitor targeting SARS-CoV-2 M^{Pro} with favorable inhibition potency and drug-like properties, thus making it a promising drug candidate for the treatment of COVID-19. We previously solved the structure of PF-07304814 in complex with SARS-CoV-2 M^{Pro}. However, the binding modes of PF-07304814 with M^{Pro}s from evolving SARS-CoV-2 variants is under-determined. In the current study, we expressed six M^{Pro} mutants (G15S, K90R, M49I, S46F, V186F, and Y54C) that have been identified in Omicron variants including the recently emerged XBB.1.16 subvariant and solved the crystal structures of PF-07304814 bound to M^{Pro} mutants. Structural analysis provided insight into the key molecular determinants responsible for the interaction between PF-07304814 and these mutant M^{Pro}s. Patterns for PF-07304814 to bind with these investigated M^{Pro} mutants and the wild-type M^{Pro} are generally similar but with some differences as revealed by detailed structural comparison. Structural insights presented in this study will inform the development of novel drugs against SARS-CoV-2 and the possible conformation changes of M^{Pro} mutants when bound to an inhibitor.

Keywords SARS-CoV-2, Variant, Main protease, Inhibitor, PF-07304814

Introduction

In late December of 2019, a novel coronavirus was detected in Wuhan city, situated in Hubei Province of central China [1, 2]. In February 2020, severe acute

respiratory syndrome coronavirus 2 (SARS-CoV-2) was designated as the name of this newly emerging coronavirus by the International Committee on Taxonomy of Viruses (ICTV). SARS-CoV-2 is known as the pathogenic agent responsible for the coronavirus disease-19 (COVID-19) [3]. Being highly transmissible, SARS-CoV-2 has spread quickly to the entire world and posed a serious threat to global public health [4, 5]. As of April 26, 2023, over 764 million confirmed cases and more than 6.9 million deaths have been reported worldwide, with the numbers still increasing (<https://covid19.who.int/>). The impressive commitment made by the biomedical research community led to rapid development of several safe and effective vaccines [6–9], which plays an important role in reducing the rates of infection, hospitalization, and mortality. However, the continuing emergence

[†]Haihai Jiang and Xiaofang Zou contributed equally to this work.

*Correspondence:

Jin Zhang
zhangxiaokong@hotmail.com
Jian Li

rmsl_2040@163.com

¹ School of Basic Medical Sciences, Nanchang University, Nanchang 330031, China

² College of Pharmaceutical Sciences, Gannan Medical University, Ganzhou 341000, China

³ Shenzhen Crystal Biopharmaceutical Co., Ltd., Shenzhen 518118, China

⁴ Jiangxi Jmerry Biopharmaceutical Co., Ltd., Ganzhou 341000, China

of SARS-CoV-2 variants highlights that the battle against COVID-19 is far from over. Five variants have been characterized by the World Health Organization (WHO) as variants of concern (VOCs), namely Alpha (B.1.1.7), Beta (B.1.351), Gamma (P.1), Delta (B.1.617.2), and Omicron (B.1.1.529), and several variants have been designated as variants of interest (VOIs), including Lambda (C.37) and Kappa (B.1.617.1). Omicron is currently the most prevalent SARS-CoV-2 variant and has evolved into many sub-variants with high immune evasion ability, including BA.5, BF.7, BQ.1, XBB.1.5, and the newly emerged XBB.1.16 [10–14], which may associate with the reduced effectiveness of available vaccines. Therefore, effective and curative treatment measures are still urgently needed to combat COVID-19.

Development of small molecules that exert antiviral efficacies against SARS-CoV-2 infection by inhibiting its main protease (M^{pro}), a vital enzyme in the viral life cycle, continues to be a relevant trend in searching for COVID-19 treatments [15]. M^{pro} is able to cleave the bio-synthesized viral polyprotein 1ab (pp1ab) at 11 specific sites to produce a variety of mature non-structural proteins necessary for viral replication and transcription [16, 17]. In most cases, this enzyme has a recognition sequence of (Leu-Gln)-(Ser-Ala-Gly), with the bond between glutamine and serine being the cleavage site [18, 19]. Such an enzymatic cleavage specificity is absent in humans, which indicates the absence of toxicity of potential M^{pro} inhibitors [18, 19]. Thus, inhibitors targeting the main protease can effectively impede the infection of SARS-CoV-2 and represent promising antiviral drug candidates.

In the last three years, various M^{pro} inhibitors have been discovered by using the high-throughput screen or structure-based drug design [20]. Among these, PF-07304814, also named lufotrelvir, is a highly soluble phosphate prodrug of PF-00835231 and a first in class 3CL inhibitor to treat SARS-CoV-2 infection [21]. It is developed by Pfizer and need to be administered by the intravenous route. When administrated, PF-07304814 will be cleaved by ubiquitous human alkaline phosphatase and metabolized into its active form, namely PF-00835231 (Fig. 1), which exhibits potent antiviral efficacy against SARS-CoV-2

as well as favorable absorption, distribution, metabolism, excretion (ADME) profiles [21]. The crystal structure of SARS-CoV-2 M^{pro} bound to PF-00835231 has been determined [22]. Recently, our group also solved the crystal structure of SARS-CoV-2 M^{pro} bound to the drug candidate PF-07304814 (PDB ID 7VVP) and illustrated the molecular mechanism for interaction [23]. The carbonyl carbons of hydroxymethyl ketone (HMK) warheads of PF-00835231 and PF-07304814 form irreversible covalent bonds with the sulfur atom of M^{pro} active-site cysteine (Cys145). These data provide structural basis and support that PF-00835231 as well as PF-07304814 is a valuable scaffold for developing effective drugs to treat COVID-19.

To date, multiple mutations in the main proteases of emerging SARS-CoV-2 variants are identified [24, 25], which may perturb the interaction network between PF-00835231 and SARS-CoV-2 M^{pro} , thus affecting its potency. PF-07304814 has the similar molecular structure with its active form. Determining the interacting details between PF-07304814 and M^{pro} mutants will provide valuable information on drug resistance to PF-00835231 and therapeutic implication for COVID-19. In this study, we solved the crystal structures of PF-07304814 in complex with several M^{pro} mutants, each carrying a previously reported single amino acid substitution, and revealed the structural basis for their interactions. The results provide structural insights for understanding the possible interaction differences between PF-07304814 and various mutant M^{pro} s, and will add to develop more effective drugs to treat viral infection caused by SARS-CoV-2 as well as its variants.

Results

Structure determination of M^{pro} mutants-PF-07304814 complexes

A total of six previously reported mutations (G15S, K90R, M49I, S46E, V186F, and Y54C) are included in this investigation. The occurrence rates for G15S, K90R, M49I, S46E, and V186F mutations among 15,476,050 sequenced SARS-CoV-2 M^{pro} genes in the Global Initiative on Sharing All Influenza Data (GISAID) (accessed

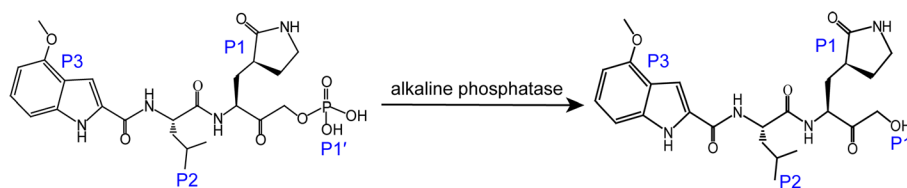


Fig. 1 SARS-CoV-2 M^{pro} inhibitor PF-07304814. PF-07304814 is a phosphate ester prodrug of PF-00835231 that is rapidly metabolized into PF-00835231 by alkaline phosphatase, The subsites of PF-07304814 and PF-00835231 are indicated

on April 27, 2023) are 0.19%, 1.34%, 0.015%, 0.025%, and 0.016%, respectively, while only 13 reports are available regarding the occurrence of Y54C mutation. All these six mutations are identified in the wildly contagious Omicron variants. G15S and K90R mutations, previously dominant in Lambda and Gamma variants, respectively, are also found in the recently emerged XBB.1.16 subvariant. In order to figure out the interaction details between PF-07304814 and these six mutant M^{Pro}s, the co-crystallization method was used to determine the structures of M^{Pro} mutants bound to PF-07304814. All M^{Pro} mutants were expressed in *E. Coli* cells, purified to near homogeneity, and then incubated with excessive PF-07304814 (Fig. 1). The prepared samples were successfully crystallized and the crystal structures of PF-07304814 in complex with these M^{Pro} mutants were solved to 2.29-Å (G15S), 1.95-Å (S46F), 2.05-Å (M49I), 1.75-Å (Y54C), 1.97-Å (K90R), and 1.70-Å (V186F) resolution, respectively. All these complex structures are in space group P12₁1, which is different with that (P2₁2₁2₁) of wild-type

M^{Pro}-PF07304814 complex [23]. Data collection and refinement statistics are summarized in Table 1.

Overall structure of M^{Pro} mutants bound to PF-07304814

In these solved complex structures, each mutant M^{Pro} is displayed as a dimer, which is exactly the form with enzymatic activity. Like the wild-type SARS-CoV-2 M^{Pro}, each protomer of the dimeric M^{Pro} mutants contains three major domains (domain I-III). Domain I resides between residues 10 through 99 and forms an anti-parallel β sheet. Domain II is located between residues 100 through 184 and is composed of anti-parallel β sheet structures. Domain III starts at the residue 201 and ends at the residue 303, which is consisting of several α helices and connects to domain II by a long loop region (residues 185–200). All the SARS-CoV-2 M^{Pro} mutants bind with two PF-07304814 molecules. PF-07304814 was found to insert into the active sites of mutant M^{Pro}s (Fig. 2), which is situated in the cleft between domains I and II. We then extracted the electron density maps of

Table 1 Data collection and refinement statistics

	SARS-CoV-2 M ^{Pro} G15S-PF-07304814	SARS-CoV-2 M ^{Pro} S46F-PF-07304814	SARS-CoV-2 M ^{Pro} M49I-PF-07304814	SARS-CoV-2 M ^{Pro} Y54C-PF-07304814	SARS-CoV-2 M ^{Pro} K90R-PF-07304814	SARS-CoV-2 M ^{Pro} V186F-PF-07304814
PDB Code	8HVU	8HVV	8HVW	8HVX	8HVV	8HVZ
Data collection						
Space group	P12 ₁ 1	P12 ₁ 1	P12 ₁ 1	P12 ₁ 1	P12 ₁ 1	P12 ₁ 1
a, b, c(Å)	55.70, 99.26, 59.9	55.29, 98.65, 59.50	55.13, 99.23, 59.64	55.16, 99.35, 59.80	54.69, 98.56, 58.86	55.33, 99.15, 59.54
α, β, γ(°)	90.00, 108.26, 90.00	90.00, 108.18, 90.00	90.00, 107.95, 90.00	90.00, 107.51, 90.00	90.00, 107.50, 90.00	90.00, 108.00, 90.00
Wavelength(Å)	0.97918	0.97918	0.97918	0.97918	0.97918	0.97918
Resolution(Å) ^a	2.29 (2.35–2.2)	1.95 (2.07–1.95)	2.05 (2.10–2.05)	1.75 (1.85–1.75)	1.97 (2.07–1.97)	1.70 (1.79–1.70)
Total reflections	182,238	207,373	225,654	371,530	256,472	386,686
Unique reflections ^a	27,534	41,127	37,920	61,385	38,245	58,251
R _{merge} (%) ^a	10.4 (85.6)	5.8 (34.0)	7.0 (94.6)	4.5 (59.1)	6.3 (67.6)	4.2 (62.5)
Mean I/σ(I) ^a	9.6/2.0	15.3/4.4	14.4/2.0	14.5/2.8	16.2/2.7	21.0/2.6
Completeness(%) ^a	98.6(99.9)	93.5(99.3)	99.0(97.1)	99.7(99.6)	90.4(86.8)	86.2(97.0)
Redundancy	6.6(6.5)	5.0(4.3)	6.0(4.3)	6.1(5.9)	6.7(6.7)	6.6(6.0)
Refinement						
Resolution (Å)	56.88–2.29	35.96–1.95	49.62–2.05	49.67–1.75	48.78–1.97	52.62–1.70
R _{work} /R _{free} ^b	21.20/25.51	19.95/23.87	21.06/24.81	21.59/23.64	21.55/24.83	21.49/23.84
Atoms	4616	4657	4601	4377	4589	4575
Mean temperature factor (Å ²)	40.6	34.9	39.5	34.4	42.8	34.2
Bond lengths (Å)	0.009	0.007	0.007	0.007	0.007	0.006
Bond angles (°)	1.156	0.873	0.861	0.911	0.852	0.864
Ramachandran plot						
Favored (%)	97.29	97.96	97.81	98.05	97.64	98.12
Allowed (%)	2.71	2.04	2.19	1.95	2.36	1.88
Outliers (%)	0	0	0	0	0	0
Molprobability	1.58	1.25	1.10	1.10	1.15	1.06

R_{merge} = $\sum_{hkl} \sum_i |I_i(hkl) - \bar{I}(hkl)| / \sum_{hkl} \bar{I}(hkl)$, where $\bar{I}(hkl)$ is the mean intensity of a set of equivalent reflections

R_{work} = $\sum_{hkl} ||F_{obs}| - |F_{calc}|| / \sum_{hkl} |F_{obs}|$, where F_{obs} and F_{calc} are observed and calculated structure factors, respectively

^a The values in parentheses are for the outermost shell

^b R_{free} is the R_{work} based on 5% of the data excluded from the refinement

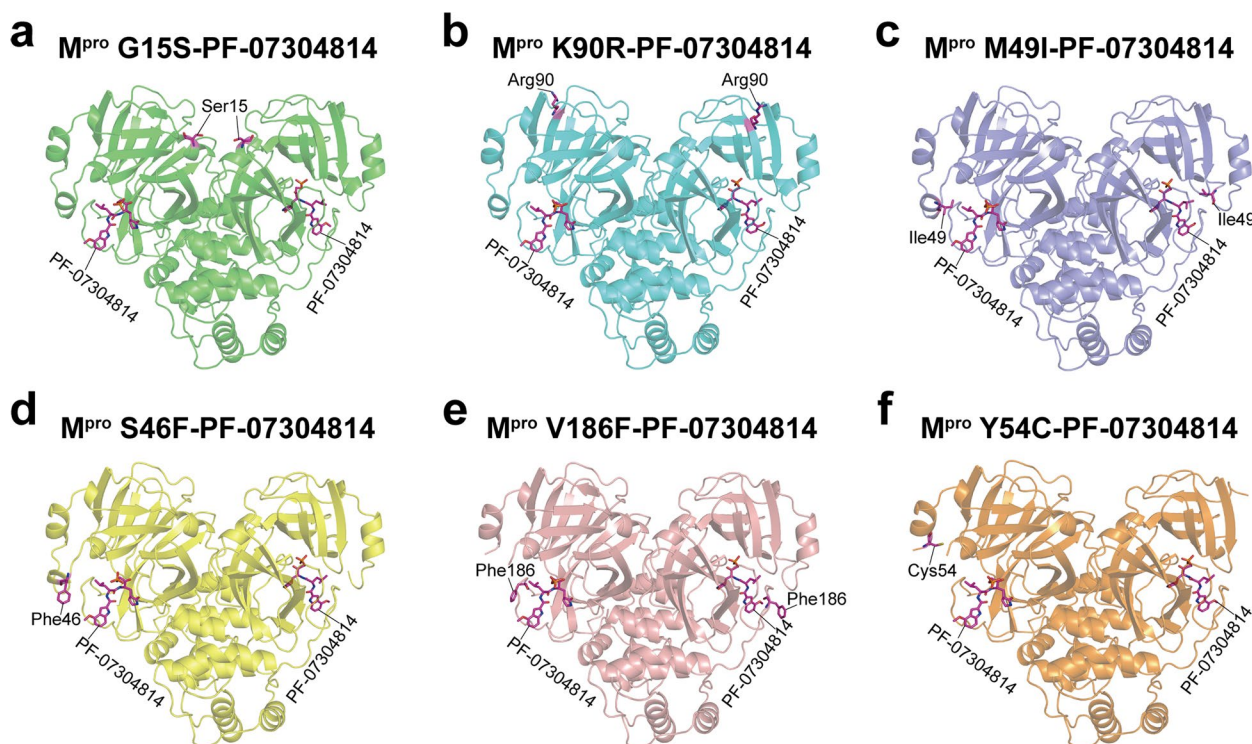


Fig. 2 Structural overview of SARS-CoV-2 M^{Pro} mutants in complex with PF-07304814. **a** SARS-CoV-2 M^{Pro} G15S (green) in complex with PF-07304814 (magenta). **b** SARS-CoV-2 M^{Pro} K90R (cyan) in complex with PF-07304814 (magenta). **c** PF-07304814 (magenta) bound to SARS-CoV-2 M^{Pro} M49I (slateblue). **d** PF-07304814 (magenta) in complex with SARS-CoV-2 M^{Pro} S46F (yellow). **e** PF-07304814 (magenta) bound to SARS-CoV-2 M^{Pro} V186F (salmon). **f** PF-07304814 (magenta) bound to SARS-CoV-2 M^{Pro} Y54C (orange). The SARS-CoV-2 M^{Pro} mutants are shown as cartoons, while PF-07304814 molecules and mutated residues are displayed as sticks

the ligand PF-07304814, the catalytic residues (including His41 and Cys145), and the mutant residues. Cys145 is a nucleophilic reagent in the hydrolysis process of SARS-CoV-2 M^{Pro} mutant, and the binding of PF-07304814 will inhibit its hydrolysis reaction. As unambiguously shown by the electron density maps, the carbonyl carbon of the hydroxymethylketone (HMK) warhead of PF-07304814 covalently interact with the sulfur atom of Cys145 in SARS-CoV-2 M^{Pro} mutants (Fig. 3). Two of these mutations (G15S and K90R) are distal to the substrate binding site of SARS-CoV-2 M^{Pro}, while the other four mutations (M49I, S46F, V186F, and Y54C) are located in or nearby the substrate binding site (Figs. 2 and 3). However, the electron densities for mutant residues S46F and Y54C could only be traced in one protomer of the M^{Pro} (Fig. 2). The quality of the fitting of two PF-07304814 molecules in the electron density maps has been evaluated by two parameters, namely the real-space R-factor (RSR) and the real-space correlation coefficient (RSCC). Calculated RSR values for the ligand 1 and ligand 2 range from 0.10 to 0.15 and from 0.08 to 0.17, respectively, while RSCC values for the ligand 1 and ligand 2 range from 0.87 to 0.93 and from 0.90 to 0.93, respectively (Table 2).

The binding energy between PF-07304814 and M^{Pro} mutants

A molecular docking analysis was performed to characterize the binding energy between PF-07304814 and the six SARS-CoV-2 M^{Pro} mutants as well as the wild-type SARS-CoV-2 M^{Pro}. The binding energy values in kcal/mol are described in Table 3. As shown in Table 3, the binding energy between wild-type M^{Pro} and PF-07304814 is -9.2 kcal/mol. In comparison, PF-07304814 shows higher binding energy with the six M^{Pro} mutants. Among these, PF-07304814 shows the least binding energy with V186F M^{Pro}, but shows the highest binding energy with M49I M^{Pro}. The binding energy values between PF-07304814 and these M^{Pro} mutants are -8.9 (V186F), -7.4 (G15S), -7.4 (K90R), -6.8 (S46F), -6.6 (Y54C), and -6.3 (M49I) kcal/mol. These interaction energy data may reflect the strength of the enzyme-ligand complex.

Comparison of binding modes of PF-07304814 with different M^{Pro} mutants

To determine whether conformation of different M^{Pro} mutants is changed when bound to PF-07304814, we superposed these six structures with wild-type

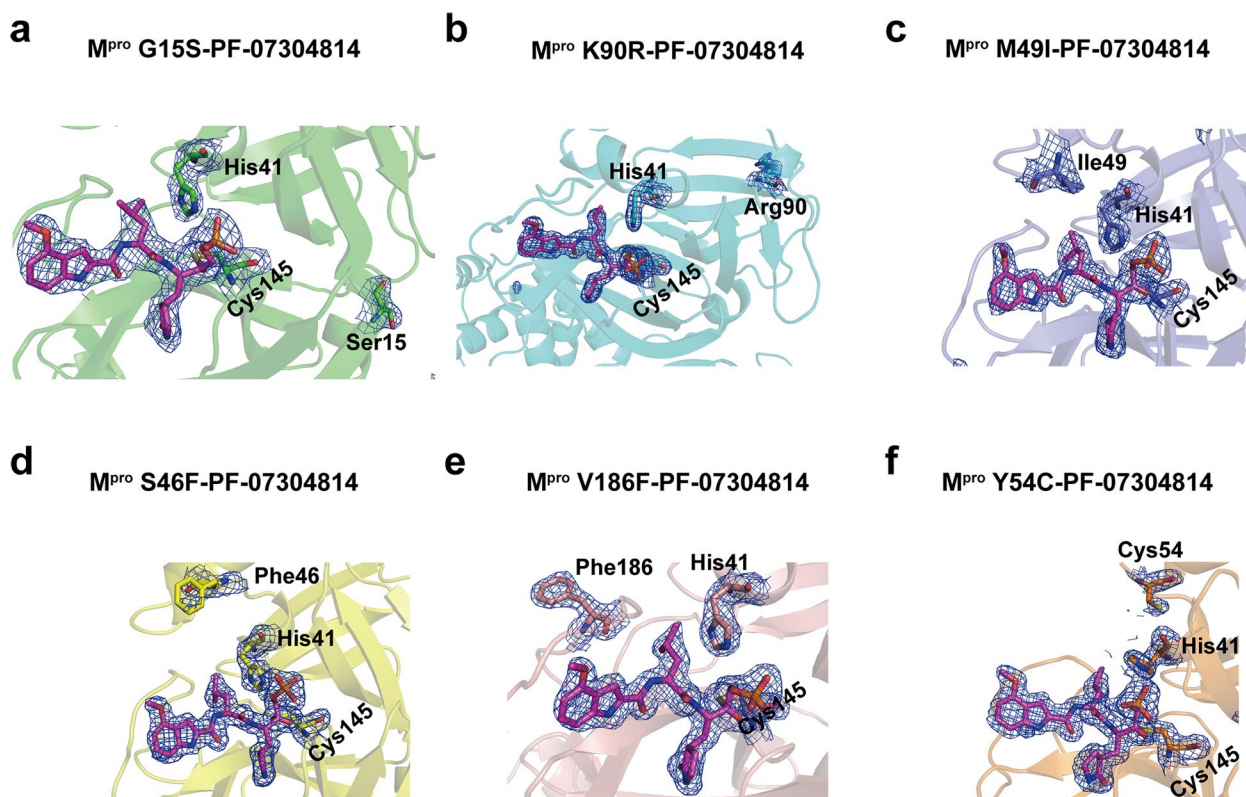


Fig. 3 The electron density maps of ligands, catalytic dyad residues, and mutant residues in M^{Pro} mutants-PF-07304814 complexes. **a-f** The 2Fo-Fc density map (1.0σ) of PF-07304814 (magentas) molecules bound to M^{Pro} G15S (**a**, green), K90R (**b**, cyan), M49I (**c**, slateblue), S46F (**d**, yellow), V186F (**e**, salmon), and Y54C (**f**, orange) are shown as a blue mesh. The ligands, catalytic dyad residues, and mutant residues are shown as sticks

Table 2 RSR and RSCC values for ligands in the SARS-CoV-2 M^{Pro} mutants-PF-07304814 complexes

Complexes	RSR (real-space R-factor)		RSCC (real-space correlation coefficient)	
	Ligand 1	Ligand 2	Ligand 1	Ligand 2
G15S-PF-07304814	0.15	0.17	0.92	0.91
K90R-PF-07304814	0.14	0.12	0.90	0.90
M49I-PF-07304814	0.13	0.12	0.91	0.92
S46F-PF-07304814	0.13	0.11	0.92	0.92
V186F-PF-07304814	0.10	0.08	0.93	0.93
Y54C-PF-07304814	0.14	0.10	0.87	0.93

M^{Pro}-PF-07304814 complex. The root mean square deviation (RMSD) values of equivalent Cα positions range from 0.695 to 0.808 Å. The results showed that the binding patterns of PF-07304814 are not significantly perturbed by these substitutions on SARS-CoV-2 M^{Pro} (Fig. 4). Indeed, two substitutions, namely G15S and K90R, are far from the binding site of PF-07304814 and have little impact on the interaction between PF-07304814 and M^{Pro}. Though four other substitutions, namely S46F, M49I, Y54C, and V186F, situated nearby the binding site of PF-07304814, no significant changes were found in the orientation of PF-07304814.

Detailed interaction between PF-07304814 and different M^{Pro} mutants

We further analyzed the interaction details between PF-07304814 and different M^{Pro} mutants. PF-07304814 is composed of four moieties, including P1', P1, P2, and

Table 3 The binding energies between PF-07304814 and SARS-CoV-2 M^{Pro} mutants

M ^{Pro} s	wild-type	G15S	K90R	M49I	S46F	V186F	Y54C
Binding energy with ligand (kcal/mol)	-9.2	-7.4	-7.4	-6.3	-6.8	-8.9	-6.6

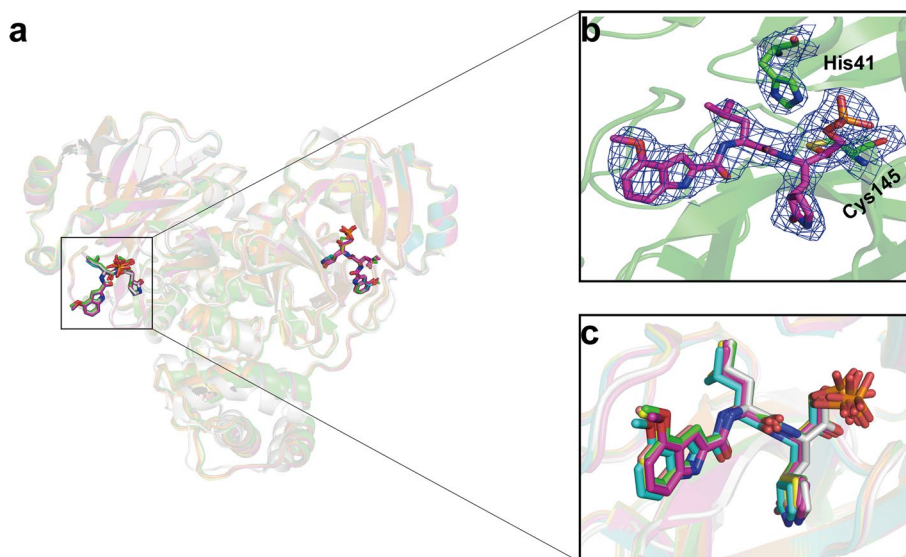


Fig. 4 Structural comparison of SARS-CoV-2 M^{Pro} -PF-07304814 complexes. **a** Overview of structural superposition of wild-type M^{Pro} -PF-07304814 complex (gray) and M^{Pro} G15S-PF-07304814 complex (green), M^{Pro} K90R-PF-07304814 complex (cyan), M^{Pro} M49I-PF-07304814 complex (slateblue), M^{Pro} S46F-PF-07304814 complex (yellow), M^{Pro} V186F-PF-07304814 complex (salmon), and M^{Pro} Y54C-PF-07304814 complex (orange). The SARS-CoV-2 M^{Pro} and its mutants are shown as cartoons, while PF-07304814 molecules are displayed as sticks. **b** The representative electronic density map of the ligands. **c** A enlarged view of structural superpositions

P3 (Fig. 1), and each substituent occupies S1', S1, S2, and S3 pockets of SARS-CoV-2 M^{Pro} , respectively. Similar with that in wild-type M^{Pro} -PF-07304814 interface [23], a C-S covalent bond is formed between the HMK warhead of PF-07304814 and the active-site cysteine (Cys145) in M^{Pro} mutants and a tetrahedral carbinol complex is generated. In addition, multiple hydrogen-bonding interactions are formed between PF-07304814 and the residues in the active site of M^{Pro} mutants (Fig. 5). A phosphate moiety is presented at the P1' site of PF-07304814. One of the hydroxyl groups forms a hydrogen bond with the Gly143 backbone NH of mutant M^{Pro} s, while the carbinol hydroxyl forms an additional hydrogen-bonding interaction with the backbone NH of Cys145 (Fig. 5). These interactions are also observed in the wild-type M^{Pro} -PF-07304814 interface. However, another hydroxyl group of the phosphate moiety of PF-07304814 forms a water molecule mediated hydrogen bond with Thr26 when bound to wild-type M^{Pro} [20]. Such an interaction can not be seen in the M^{Pro} mutants-PF-07304814 complexes. The S1 pocket of SARS-CoV-2 M^{Pro} has a strong preference for Gln at the P1 position. Upon binding to the wild-type M^{Pro} or mutant M^{Pro} s, the lactam ring displayed at the P1 position of PF-07304814 forms hydrogen-bonding interactions with the N ϵ 2 of H163, the backbone oxygen of Phe140, and the side-chain oxygen of Glu166. This moiety of the inhibitor is expected to mimic Gln at the P1 position of the substrate. Moreover, a hydrogen-bonding interaction can be seen between the P1 NH in PF-07304814 and the main-chain carbonyl

oxygen of His164 in wild-type and mutant M^{Pro} s. A leucine moiety is presented at the P2 position of PF-07304814, while an indole group is displayed at the P3 position. The P2 NH forms a hydrogen-bonding interaction with the side chain of Gln189 in wild-type and mutant M^{Pro} s (except Y54C). In M^{Pro} Y54C-PF-07304814 complex, the hydrogen bond between the Gln189 of mutant M^{Pro} and the inhibitor can not be observed. For both wild-type and mutant M^{Pro} s, the P3 indole of PF-07304814 forms two another hydrogen-bonding interactions with the backbone carbonyl and NH of Glu166. These structural details established that how lufotrelvir recognizes different SARS-CoV-2 M^{Pro} mutants, which will be helpful in drug resistance monitoring and further drug development.

Discussion

Effective antiviral drugs are urgently required to combat the public health threat caused by SARS-CoV-2 and its variants. Main protease (M^{Pro}) serves as a superior target for anti-SARS-CoV-2 drugs. PF-07304814, also named lufotrelvir, is one of the most advanced M^{Pro} inhibitors that have entered clinical trials and also the phosphate prodrug of PF-00835231 that could increase its bioavailability [21]. In addition, PF-07304814 exhibits good tolerability, ADME, and safety in rat models and preclinical trials [21]. Even that the phase 1 data of PF-07304814 were not released and the phase 2/3 trial of PF-07304814 was also suspended, such a promising M^{Pro} inhibitor provides a valuable scaffold for drug design. In fact,

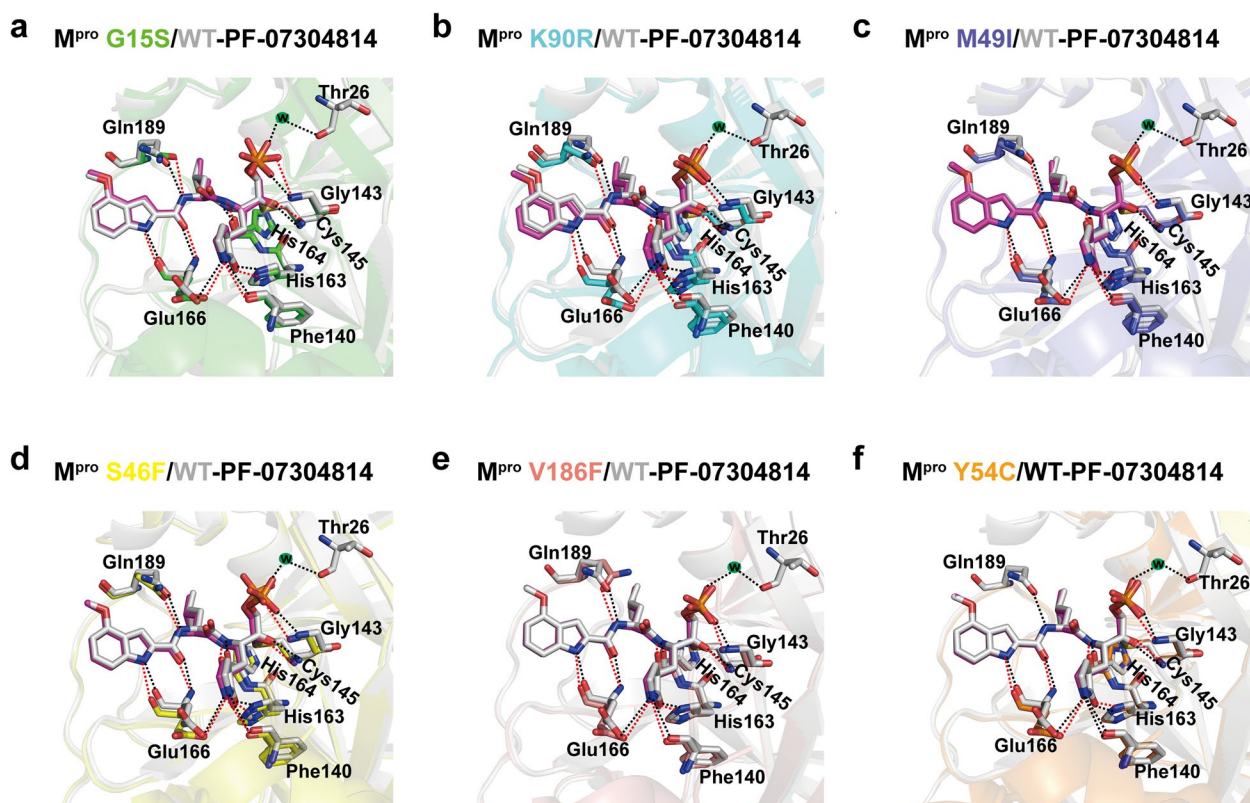


Fig. 5 Structural characterization of PF-07304814 bound to SARS-CoV-2 wild-type M^{pro} or M^{pro} mutants. **a-f** Superposition of the x-ray crystal structures of PF-07304814 bound to wild-type (WT) SARS-CoV-2 M^{pro} (**a-f**, gray) and SARS-CoV-2 M^{pro} G15S (**a**, green), K90R (**b**, cyan), M49I (**c**, slateblue), S46F (**d**, yellow), V186F (**e**, salmon), and Y54C (**f**, orange). PF-07304814 molecules bound to WT M^{pro} are displayed as gray sticks, while PF-07304814 molecules bound to mutant M^{pro}s are shown as magenta sticks. Hydrogen bonds are shown as black dashed lines between WT M^{pro} protein and the inhibitor, while hydrogen bonds are shown as red dashed lines in M^{pro} mutants-PF-07304814 complexes

PF-07321332 (nirmatrelvir), the active ingredient in the oral COVID-19 drug PAXLOVID™, was developed by the modification of PF-07304814. Moreover, as a phosphate ester prodrug of PF-00835231, PF-07304814 displays a very similar molecule structure with its active form. This study used structural biology methods to reveal the differences and similarities of PF-07304814 in binding with different M^{pro} mutants and evaluated the possible conformation change of mutant M^{pro}s when bound to this inhibitor. Information on structural changes of M^{pro} mutants when bound to PF-07304814 can be useful for understanding how to inhibit mutant M^{pro}s with potential drug resistance and for optimizing the inhibition potency of PF-07321332 as well as PF-00835231. Thus, these data will be informative to develop next generation drugs against wild-type SARS-CoV-2 and its variants.

The use of the electronic density to measure the model fit has some limitation in catching all possible problems in the model [26]. A good manner to evaluate how well a subset of atomic coordinates fits the experimental electron density is the Real Space R-factor (RSR) [27].

Another well-established measure of model fit to the experimental data is the real-space correlation coefficient (RSCC) [28]. Moreover, a RSR value less than 0.3 indicates a good agreement between observed and calculated electron densities and a RSCC value greater than 0.8 indicates a good correlation between calculated and observed electron density. In the present study, the RSR and RSCC values for the ligand PF-07304814 range from 0.08 to 0.17 and from 0.87 to 0.93, respectively, thus indicating a good consistency between the model and the density map for the co-crystallized ligands.

As determined in this study, the binding pattern of PF-07304814 is not significantly perturbed by the six single mutations (G15S, K90R, M49I, S46F, V186F, and Y54C), and the ligand largely maintaining the interaction networks observed in the wild-type M^{pro} with only slight differences. One obvious difference is that a water-mediated hydrogen bond with Thr26 can not be found in the interfaces of mutant M^{pro}-PF-07304814 complexes. Another obvious difference is that the hydrogen bond between Gln189 from M^{pro} Y54C and PF-07304814 can

not be found. A water bridge between the ligand and M^{PRO} Thr26 plays an important role in ligand binding [29]. Gln189 of SARS-CoV-2 M^{PRO} has also been identified as one of the hot spot residues in the interaction with inhibitors and a hydrogen bonding with Gln189 may enhance the potency of M^{PRO} inhibitor [29, 30]. These differences may impact the binding affinities between M^{PRO} mutants and PF-07304814.

We then calculate the binding energies between PF-07304814 and the targeted protein to have a glimpse of the potential of PF-07304814 as a ligand to the active site of M^{PRO} mutants. Specially, the binding between PF-07304814 and Y54C mutant is much higher than that between PF-07304814 and wild-type M^{PRO}, which may be due to the lost of hydrogen bonding interaction between the Gln189 in M^{PRO} Y54C and the ligand PF-07304814. Other M^{PRO} mutants also show higher binding energy values with PF-07304814 compared to wild-type M^{PRO}, which may be due to the lost of a water mediated hydrogen bond between Thr26 of M^{PRO}s. As evidenced in the literature, the covalent docking may not be able to accurately reveal the real binding strength between the ligand and targeted proteins [31], because covalent docking disregards to explicitly explore the reactivity of the covalent inhibitors. Thus, further biochemical assays are needed in the future to evaluate the inhibition efficacy of PF-07304814 against these M^{PRO} mutants.

The present study investigated the impact of single amino acid substitutions of M^{PRO}s on the its interaction with PF-07304814. However, main proteases from the circulating SARS-CoV-2 strains tend to possess multiple mutations. Next steps should also include experimental and structural characterization of the impact of combined mutations on the interactions network between PF-07304814 and SARS-CoV-2 M^{PRO}. As PF-07304814 is a promising drug candidate to treat COVID-19, M^{PRO} mutations that cover a larger variety of SARS-CoV-2 lineages should also be investigated in further study.

Materials and methods

Protein expression and purification

Expression of M^{PRO} mutants of SARS-CoV-2 (including G15S, S46F, M49I, Y54C, K90R, and V186F) was performed according to previous description [23]. Briefly, pET-28a plasmids containing the entire coding sequences were transformed into competent *Escherichia coli* (*E. coli*) Rosetta DE3 cells. After grown in LB (Luria-Bertani) broth with the OD₆₀₀ reaching 0.6–0.8 and induced with 500 μmol/L isopropyl-β-d-thiogalactopyranoside (IPTG), the *E. coli* cells started to produce mutant M^{PRO} proteins. The cells were then centrifuged at 10,000 × g for 15 min at 4 °C. The supernatants were discarded, while the cell pellets were collected, resuspended using lysis buffer, and

lysed by sonication. The cleared lysate was subjected to immobilized metal affinity chromatography on a HisTrap HP 5 mL column (Cytiva). Imidazole gradient treatment was employed to elute the target proteins. Elution fractions containing mutant M^{PRO} proteins were harvested and incubated with TEV protease to remove the N-terminal His-tag. The digested M^{PRO}s were subjected to gel filtration for further purification. High-purity mutant M^{PRO} fractions were collected with centrifuge tubes and concentrated using concentrator tubes with the molecular weight cutoff of 10 kDa (Millipore).

Crystallization

The eluted SARS-CoV-2 M^{PRO} mutants were concentrated to 5 mg/mL, and then incubated with PF-07304814 molecules at a molar ratio of 1:3 on ice for 30 min. Crystallization was carried out at 18 °C via the sitting drop vapor diffusion technique. After several days, the crystals of six SARS-CoV-2 M^{PRO} mutants in complex with PF-07304814 were obtained. The final crystallization condition of the M^{PRO} G15S-PF-07304814 complex was 0.15 M Tris pH 8.0, 30% (w/v) PEG 4000. The final crystallization condition of the M^{PRO} S46F-07304814 complex was 0.20 M Na₂SO₄, 24% (w/v) PEG 3350. The final crystallization condition of the M^{PRO} M49I-07304814 complex was 0.15 M Na₂SO₄, 20% (w/v) PEG 3350. The final crystallization condition of the M^{PRO} Y54C-07304814 complex was 0.1 M BIS-Tris pH 6.5, 20% (w/v) PEG 3350. The final crystallization condition of the M^{PRO} K90R-07304814 complex was 0.15 M Na₂SO₄, 22% (w/v) PEG 3350. The final crystallization condition of the M^{PRO} V186F-07304814 complex was 0.16 M Na₂SO₄, 20% (w/v) PEG 3350.

Data Collection, structure determination, and refinement

Before X-ray diffraction data collection, the crystals of SARS-CoV-2 M^{PRO} mutants in complex with PF-07304814 were soaked in the crystallization buffer consisting of 20% glycerol and then flash cooled in liquid nitrogen. X-ray diffraction data were collected at 100 K at BL10U2 of the Shanghai Synchrotron Radiation Facility (SSRF). All the data sets were processed using HKL2000 software [32]. All the complex structures were determined via the molecular replacement method by using the Phaser program [33]. Coot and Phenix softwares were used for atomic model building and maximum likelihood-based refinement [34, 35]. The stereochemical qualities of the final models were assessed with MolProbity [36]. The data collection and processing statistics and structural refinement statistics for the M^{PRO} mutants-PF-07304814 complexes are shown in Table 1. The values of real-space R-factor (RSR) and real-space correlation

coefficient (RSCC) were calculated to measure the electron density fit for PF-07304814 molecules in the complex structures [27, 28].

Molecular docking analysis

The co-crystallized M^{Pro} mutants and the ligands were prepared in advance of docking. For M^{Pro} mutants, the chain B of protein was kept and other atoms were deleted. For the ligand PF-07304814, the residue Cys145 (atoms SG, CA, CB, and C) where Michael addition reaction took place was appended to guide covalent docking. Then, the AutoDockTools-1.5.7 program was employed to add polar hydrogen atoms and assign Kollman charges. The grid box defining the binding pocket was set as the center of the ligand with length of 30 angstrom to all three dimensions. Finally, the ADFR program was applied to generate the affinity maps and perform covalent docking [37].

Acknowledgements

H.J. was supported by the Natural Science Foundation of Jiangxi Province (20224BAB216004). J.L. was supported by Jiangxi Natural Science Foundation for Distinguished Young Scholars (20212ACB216001), Jiangxi Key Research and Development Program (20203BBG73063), Jiangxi “Double Thousand Plan (jxsq2019101064)”, and the foundation of Gannan Medical University (QD201910). J.Z. was supported by the National Natural Science Foundation of China (32271260), the CAS “Light of West China” Program (xbzg-zdsys-202005), and the Natural Science Foundation of Jiangxi Province (20224ACB206046).

Authors' contributions

JL and JZ designed experiments. HJ, XFZ, and PZ purified and crystallized the protein as well as performed the related assays. PZ, XYZ, XLZ, and JW collected the diffraction data. HJ and XFZ prepared the figures. HJ wrote the manuscript. HJ, JZ, and JL revised the manuscript. The authors read and approved the final manuscript.

Availability of data and materials

Coordinates and structure factors for SARS-CoV-2 M^{Pro} G15S, S46F, and M49I in complex with PF-07304814 have been deposited in PDB under accession code 8HVU, 8HVV, and 8HVV, respectively. Coordinates and structure factors for SARS-CoV-2 M^{Pro} Y54C, K90R, and V186F in complex with PF-07304814 have been deposited in PDB under accession code 8HVX, 8HVV, and 8HVZ, respectively.

Declarations

Ethics approval and consent to participate

Not applicable.

Competing interests

The authors declare that no conflicts of interest exist.

Received: 4 February 2023 Accepted: 8 June 2023

Published online: 03 August 2023

References

- Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet*. 2020;395(10223):497–506. [https://doi.org/10.1016/S0140-6736\(20\)30183-5](https://doi.org/10.1016/S0140-6736(20)30183-5).
- Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, et al. A novel coronavirus from patients with Pneumonia in China, 2019. *N Engl J Med*. 2020;382(8):727–33. <https://doi.org/10.1056/NEJMoa2001017>.
- Coronaviridae Study Group of the International Committee on Taxonomy of V. The species severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. *Nat Microbiol*. 2020;5(4):536–44. <https://doi.org/10.1038/s41564-020-0695-z>.
- Wu JT, Leung K, Leung GM. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. *Lancet*. 2020;395(10225):689–97. [https://doi.org/10.1016/S0140-6736\(20\)30260-9](https://doi.org/10.1016/S0140-6736(20)30260-9).
- Hui DS, Madani EA, Ntoumi TA, Kock F, Dar R. The continuing 2019-nCoV epidemic threat of novel coronaviruses to global health - the latest 2019 novel coronavirus outbreak in Wuhan, China. *Int J Infect Dis*. 2020;91:264–6. <https://doi.org/10.1016/j.ijid.2020.01.009>.
- van Doremalen N, Lambe T, Spencer A, Belij-Rammerstorfer S, Purushotham JN, Port JR, et al. ChAdOx1 nCoV-19 vaccine prevents SARS-CoV-2 pneumonia in rhesus macaques. *Nature*. 2020;586(7830):578–82. <https://doi.org/10.1038/s41586-020-2608-y>.
- Baden LR, El Sahly HM, Essink B, Kotloff K, Frey S, Novak R, et al. Efficacy and safety of the mRNA-1273 SARS-CoV-2 vaccine. *N Engl J Med*. 2021;384(5):403–16. <https://doi.org/10.1056/NEJMoa2035389>.
- Polack FP, Thomas SJ, Kitchin N, Absalon J, Gurtman A, Lockhart S, et al. Safety and Efficacy of the BNT162b2 mRNA Covid-19 vaccine. *N Engl J Med*. 2020;383(27):2603–15. <https://doi.org/10.1056/NEJMoa2034577>.
- Jin L, Li Z, Zhang X, Li J, Zhu F, CoronaVac. A review of efficacy, safety, and immunogenicity of the inactivated vaccine against SARS-CoV-2. *Hum Vaccin Immunother*. 2022;18(6):2096970. <https://doi.org/10.1080/21645515.2022.2096970>.
- Kurhade C, Zou J, Xia H, Liu M, Chang HC, Ren P, et al. Low neutralization of SARS-CoV-2 Omicron BA.2.75.2, BQ.1.1 and XBB.1 by parental mRNA vaccine or a BA.5 bivalent booster. *Nat Med*. 2023;29(2):344–7. <https://doi.org/10.1038/s41591-022-02162-x>.
- Shrestha LB, Foster C, Rawlinson W, Tedla N, Bull RA. Evolution of the SARS-CoV-2 omicron variants BA.1 to BA.5: implications for immune escape and transmission. *Rev Med Virol*. 2022;32(5):e2381. <https://doi.org/10.1002/rmv.2381>.
- Qu P, Evans JP, Faraone JN, Zheng YM, Carlin C, Anghelina M, et al. Enhanced neutralization resistance of SARS-CoV-2 Omicron subvariants BQ.1, BQ.1.1, BA.4.6, BF.7, and BA.2.75.2. *Cell Host Microbe*. 2023;31(1):9–17e3. <https://doi.org/10.1016/j.chom.2022.11.012>.
- Zhu A, Wei P, Man M, Liu X, Ji T, Chen J, et al. Antigenic characterization of SARS-CoV-2 Omicron subvariants XBB.1.5, BQ.1, BQ.1.1, BF.7 and BA.2.75.2. *Signal Transduct Target Ther*. 2023;8(1):125. <https://doi.org/10.1038/s41392-023-01391-x>.
- Zappa M, Verdecchia P, Angeli F. Is the competition between variants the end of severe acute respiratory syndrome coronavirus 2 pandemic? A journey from Wuhan to XBB.1.16. *Eur J Intern Med*. 2023. <https://doi.org/10.1016/j.ejim.2023.04.016>.
- Kronenberger T, Laufer SA, Pillaiyar T. COVID-19 therapeutics: small-molecule drug development targeting SARS-CoV-2 main protease. *Drug Discov Today*. 2023;28(6):103579. <https://doi.org/10.1016/j.drudis.2023.103579>.
- Zhao Y, Zhu Y, Liu X, Jin Z, Duan Y, Zhang Q, 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(16):e2117142119. <https://doi.org/10.1073/pnas.2117142119>.
- Yadav R, Courouble VV, Dey SK, Harrison J, Timm J, Hopkins JB, et al. Biochemical and structural insights into SARS-CoV-2 polyprotein processing by Mpro. *Sci Adv*. 2022;8(49):eadd2191. <https://doi.org/10.1126/sciadv.add2191>.
- Lee J, Kenward C, Worrall LJ, Vuckovic M, Gentile F, Ton AT, et al. X-ray crystallographic characterization of the SARS-CoV-2 main protease polyprotein cleavage sites essential for viral processing and maturation. *Nat Commun*. 2022;13(1):5196. <https://doi.org/10.1038/s41467-022-32854-4>.
- Xiong M, Su H, Zhao W, Xie H, Shao Q, Xu Y. What coronavirus 3 C-like protease tells us: from structure, substrate selectivity, to inhibitor design. *Med Res Rev*. 2021;41(4):1965–98. <https://doi.org/10.1002/med.21783>.
- Agost-Beltran L, de la Hoz-Rodriguez S, Bou-Iserte L, Rodriguez S, Fernandez-de-la-Pradilla A, Gonzalez FV. Advances in the development of SARS-CoV-2 Mpro inhibitors. *Molecules*. 2022;27(8). <https://doi.org/10.3390/molecules27082523>.

21. Boras B, Jones RM, Anson BJ, Arenson D, Aschenbrenner L, Bakowski MA, et al. Preclinical characterization of an intravenous coronavirus 3CL protease inhibitor for the potential treatment of COVID-19. *Nat Commun.* 2021;12(1):6055. <https://doi.org/10.1038/s41467-021-26239-2>.
22. Hoffman RL, Kania RS, Brothers MA, Davies JF, Ferre RA, Gajiwala KS, et al. Discovery of Ketone-Based covalent inhibitors of Coronavirus 3CL Proteases for the potential therapeutic treatment of COVID-19. *J Med Chem.* 2020;63(21):12725–47. <https://doi.org/10.1021/acs.jmedchem.0c01063>.
23. Li J, Lin C, Zhou X, Zhong F, Zeng P, McCormick PJ, et al. Structural basis of main proteases of coronavirus bound to drug candidate PF-07304814. *J Mol Biol.* 2022;434(16):167706. <https://doi.org/10.1016/j.jmb.2022.167706>.
24. Motyan JA, Mahdi M, Hoffka G, Tozser J. Potential resistance of SARS-CoV-2 main protease (Mpro) against protease inhibitors: Lessons learned from HIV-1 protease. *Int J Mol Sci.* 2022;23(7). <https://doi.org/10.3390/ijms23073507>.
25. Padhi AK, Tripathi T. Hotspot residues and resistance mutations in the nirmatrelvir-binding site of SARS-CoV-2 main protease: design, identification, and correlation with globally circulating viral genomes. *Biochem Biophys Res Commun.* 2022;629:54–60. <https://doi.org/10.1016/j.bbrc.2022.09.010>.
26. Richardson JS, Richardson DC. Studying and polishing the PDB's macromolecules. *Biopolymers.* 2013;99(3):170–82. <https://doi.org/10.1002/bip.22108>.
27. Jones TA, Zou JY, Cowan SW, Kjeldgaard M. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr A.* 1991;47(Pt 2):110–9. <https://doi.org/10.1107/s0108767390010224>.
28. Tickle IJ. Statistical quality indicators for electron-density maps. *Acta Crystallogr D Biol Crystallogr.* 2012;68(Pt 4):454–67. <https://doi.org/10.1107/S0907444911035918>.
29. Thakur A, Sharma G, Badavath VN, Jayaprakash V, Merz KM Jr, Blum G, et al. Primer for Designing Main protease (M(pro)) inhibitors of SARS-CoV-2. *J Phys Chem Lett.* 2022;13(25):5776–86. <https://doi.org/10.1021/acs.jpcclett.2c01193>.
30. Chan HTH, Moesser MA, Walters RK, Malla TR, Twidale RM, John T, et al. Discovery of SARS-CoV-2 M(pro) peptide inhibitors from modelling substrate and ligand binding. *Chem Sci.* 2021;12(41):13686–703. <https://doi.org/10.1039/d1sc03628a>.
31. Sanches M, Duffy NM, Talukdar M, Thevakumaran N, Chiovitti D, Canny MD, et al. Structure and mechanism of action of the hydroxy-aryl-aldehyde class of IRE1 endoribonuclease inhibitors. *Nat Commun.* 2014;5:4202. <https://doi.org/10.1038/ncomms5202>.
32. Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* 1997;276:307–26. [https://doi.org/10.1016/S0076-6879\(97\)76066-X](https://doi.org/10.1016/S0076-6879(97)76066-X).
33. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. *J Appl Crystallogr.* 2007;40(Pt 4):658–74. <https://doi.org/10.1107/S0021889807021206>.
34. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr.* 2010;66(Pt 2):213–21. <https://doi.org/10.1107/S0907444909052925>.
35. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr.* 2010;66(Pt 4):486–501. <https://doi.org/10.1107/S0907444910007493>.
36. Chen VB, Arendall WB 3rd, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr.* 2010;66(Pt 1):12–21. <https://doi.org/10.1107/S0907444909042073>.
37. Bianco G, Forli S, Goodsell DS, Olson AJ. Covalent docking using auto-dock: two-point attractor and flexible side chain methods. *Protein Sci.* 2016;25(1):295–301. <https://doi.org/10.1002/pro.2733>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.