Peptidyl Nitroalkene Inhibitors of Main Protease (Mpro) rationalized by Computational/Crystallographic Investigations as Antivirals against SARS-CoV-2

Francisco Medrano

   Centro de Investigaciones Biológicas Margarita Salas(CIBMS), CSIC   https://orcid.org/0000-0002-8185-9751

Sergio de la Hoz-Rodríguez

   Universitat Jaume I

Sergio Martí

   Universitat Jaume I   https://orcid.org/0000-0002-1087-7143

Kemel Arafet

   Universitat Jaume I

Tanja Schirmeister

   University of Mainz

Stefan Hammerschmidt

   Johannes Gutenberg-University Mainz   https://orcid.org/0000-0002-0769-8435

Christin Müller

   University of Giessen   https://orcid.org/0000-0002-6444-9478

Águeda González-Martínez

   Centro de Investigaciones Biológicas Margarita Salas (CSIC)x

Elena Santillana Heras

   Centro de Investigaciones Biológicas (CIB), CSIC

John Ziebuhr

   Justus Liebig University Giessen   https://orcid.org/0000-0002-5741-8825

Antonio Romero Garrido

   Centro de Investigaciones Biológicas (CIB), CSIC   https://orcid.org/0000-0002-6990-6973

Alessio Lodola

   https://orcid.org/0000-0002-8675-1002

Katarzyna Swiderek

   Universitat Jaume I   https://orcid.org/0000-0002-7528-1551

Vicent Moliner

   Universitat Jaume I   https://orcid.org/0000-0002-3665-3391

Florenci González (✉ fgonzale@uji.es)

Additional Declarations: There is **NO** Competing Interest.

Peptidyl Nitroalkene Inhibitors of Main Protease (M^{pro}) rationalized by Computational/Crystallographic Investigations as Antivirals against SARS-CoV-2

Francisco J. Medrano[1]*, Sergio de la Hoz-Rodríguez[2], Sergio Martí[5], Kemel Arafet[5], Tanja Schirmeister[3], Stefan J. Hammerschmidt[3], Christin Müller[4], Águeda González-Martínez[1], Elena Santillana[1], John Ziebuhr[4], Antonio Romero[1], Alessio Lodola[6], Katarzyna Świderek[5], Vicent Moliner[5]*, Florenci V. González[2]*

[1]Centro de Investigaciones Biológicas Margarita Salas (CSIC), Ramiro de Maeztu 9, 28040 Madrid, Spain.

[2]Departament de Química Inorgànica i Orgànica, Universitat Jaume I, 12071-Castelló, Spain.

[3]Institute of Pharmaceutical and Biochemical Sciences, Johannes Gutenberg-University Mainz, Staudinger Weg 5, 55128 Mainz, Germany.

[4]Institute of Medical Virology, Justus Liebig University Giessen, Schubertstrasse 81, 35392 Giessen, Germany.

[5]Departament de Química Física i Analítica, Universitat Jaume I, 12071-Castelló, Spain.

[6]Dipartimento di Scienze degli Alimenti e del Farmaco, Università degli Studi di Parma, Italy.

Abstract

The coronavirus disease 2019 (COVID-19) pandemic continues to represent a global public health issue. The viral main protease (M^{pro}) represents one of the most attractive targets for the development of antiviral drugs. Herein we report peptidyl nitroalkenes exhibited enzyme inhibitory activity against M^{pro} (K_i: 1-10 μM) and three of them good anti-SARS-CoV-2 infection activity in the low micromolar range (EC$_{50}$: 1-12 μM) without significant toxicity. Additional kinetic studies of compounds **FGA145**, **FGA146** and **FGA147** show that all three compounds inhibit Cathepsin L, denoting a possible multitarget effect of these compounds in the antiviral activity. QM/MM computer simulations assisted in the design and in elucidating the way of action. Finally, structural analysis shows, in agreement with the computer predictions, the binding mode of **FGA146** and **FGA147** to the active site of the protein. Our results illustrate that peptidyl nitroalkenes are potent covalent reversible inhibitors of the M^{pro} and cathepsin L, and that inhibitors **FGA145**, **FGA146** and **FGA147** prevent infection becoming promising drugs against SARS-CoV-2.

Introduction

The impact of SARS-CoV-2 pandemic (COVID-19) has made the search for new therapies against coronaviruses urgent. The pandemic has resulted so far in over 600 million infections and over 6 million deaths worldwide, according to World Health Organization.[1] Recent situation in China shows the pandemic is far from over.[2] The evolution of SARS-CoV-2 virus has resulted in several highly contagious SARS-CoV2 strains that evade antibodies targeting the receptor binding domain (RBD) and threaten the effectiveness of the current vaccines.[3,4]

Among approved antivirals for the treatment of COVID-19 are nucleoside derivatives remdesivir [5] and molnupiravir,[6] with uncertain efficacy for certain types of patients, and Paxlovid,[7] a combination of M^{pro}

inhibitor nirmatrelvir and HIV protease inhibitor ritonavir. Despite of remarkable efficacy of Paxlovid, it cannot be administered to patients with liver or kidney dysfunction. Furthermore, ritonavir, which blocks the rapid metabolism of nirmatrelvir by CYP3A, interacts in turn with other drugs limiting its use. These issues demonstrate that there is an urgent need to find new antivirals for SARS-CoV-2 and for other coronavirus outbreaks in the future.

During the replication cycle the coronavirus express two overlapping polyproteins, (pp1a and pp1b) and four structural proteins from the viral RNA.[8] In order to liberate the mature viral proteins required for replication, these polyproteins must be properly processed. There are two proteases coded in the viral genome, known as 3-chymotrypsin-like protease (3CL^{pro}, 3CLP or nsp5, also termed main protease M^{pro}) and papain-like protease (PL^{pro} or nsp3). Most of the cleavages are carried out by M^{pro}.[9,10] Both M^{pro} and PL^{pro} are cysteine proteases with different site specificities.

M^{pro} is a three-domain cysteine protease essential for most maturation events within the precursor polyprotein.[9-11] The active protease is a homodimer. The active site is made up by a non-canonical Cys-His dyad located in the cleft between domains I and II.[10-12] M^{pro} hydrolyses proteins predominantly between a P1 glutamine and a small P1' amino acid, such as alanine, serine or glycine. For the P2 position, leucine is the most common amino acid in the sequence specificity for coronaviruses. There is no human homolog of M^{pro} which makes it an ideal antiviral drug target.[13-16]

Peptidyl compounds have been shown to inhibit M^{pro} in *in vitro* assays and to prevent infection by SARS-CoV-2 in cell culture.[17] The warhead of these reported inhibitors are carbonyl groups [aldehydes,[18] ketoamides,[19] ketones [20]], nitriles (as marketed inhibitor nirmatrelvir) [21] and enoates.[22] New warheads to supplement the current repertoire would be welcome and potentially more effective. In this context, the nitroalkene moiety has been previously reported by us as a valid warhead for inhibitors against cysteine proteases belonging to the papain family.[23] Irreversible inhibitors such as enoates can give rise to undesired side reactions. Alternatively, nitroalkenes represent Michael acceptor inhibitors following a reversible mode-of-action due to the low basicity of the nitronate intermediate.[24] Computational studies previously reported by us pointed to nitroalkenes as promising inhibitors of SARS-CoV-2 M^{pro}.[25] Interestingly, peptidyl nitroalkenes resulted to be potent inhibitors of human cathepsin L (CatL) as we previously reported.[24] CatL has been also recognized as a potential target for the search of drugs against COVID-19 as it is found to enable viral cell entry by activating the SARS-CoV-2 spike protein by cleavage.[26-29]

Analysis of interaction energies between the substrate (the peptide in the proteolysis reaction or the inhibitor in the case of the inhibition reaction) and the different binding pockets of SARS-CoV-2 M^{pro} based on multiscale quantum mechanics / molecular mechanics (QM/MM) studies, indicated that they are dominated by those in the P1:::S1 site.[25, 30, 31] However, the recognition portion dictates how the inhibitor is accommodated in the active site, which in turn affects the subsequent chemical reaction step. Consequently, the reactivity of the warhead and the favorable interactions between the recognition portion and the active site of the enzyme must be considered to design an efficient inhibitor.[31] In all, the experience accumulated
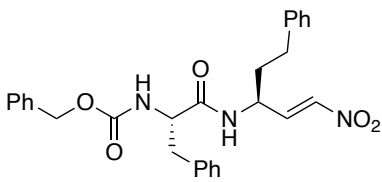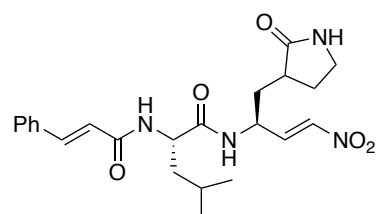
73   based on the results derived from previous studies on this and other cysteine proteases can be used to guide

74   the design of new compounds, and QM/MM simulations can be considered a useful tool to get a detailed

75   description of the chemical steps of the inhibition of protein targets by covalent inhibitors.

76   Based on our previous studies on proteolysis reaction of the SARS-CoV-2 M^{pro} [30] and its inhibition mechanisms

77   by peptidyl inhibitors with different warheads,[25, 30, 31] we have designed and synthesized six peptidyl inhibitors

78   with a nitroalkene warhead. These six new inhibitors were able to inhibit the M^{pro} *in vitro* activity in a

79   reversible mode in the low micromolar range, and three of them were found to prevent SARS-CoV-2 infection

80   in cell culture in the low micromolar range. In addition, these compounds were also tested against cathepsin

81   L, a key enzyme for the viral entry into the cells. Molecular dynamics (MD) simulations with multiscale

82   QM/MM potentials were carried out to obtain the full free energy landscape of the inhibition reaction with

83   the two most active inhibitors, confirming the interactions established with the active site residues of SARS-

84   CoV-2 M^{pro} as well as their mechanism of action for the enzyme-inhibitor covalent complexes formation.

85   Finally, the crystal structures of M^{pro} in complex with these two most active inhibitors were solved to provide

86   detailed information about the binding to SARS-CoV-2 M^{pro}.

Results

Design and synthesis of the peptidyl nitroalkenes inhibitors. Six inhibitors were designed and synthesized: three of them having the typical coronaviral protease glutamine surrogate (beta-lactam) at P1 site and a L-leucine at P2 site (**FGA145**, **FGA146** and **FGA147**), other three inhibitors display the typical cathepsin like inhibitors backbone: two having a L-homophenylalanine at P1 site and a L-leucine at P2 site (**FGA159** and **FGA177**), and one having a L-homophenylalanine at P1 site and a L-phenylalanine at P2 site (**FGA86**). For the synthesis of the ones having the glutamine surrogate at P1, the *N*-Boc protected amino alcohol **1** was firstly prepared as previously reported starting from L-glutamic acid.[32] The alcohol was then submitted to oxidation followed by a nitroaldol reaction with nitromethane in a one-pot procedure. The mixture of nitroaldols was then transformed into the corresponding inhibitors following a three-step sequence: Boc deprotection, peptide coupling and then dehydration through mesylate activation (Scheme 1). For the synthesis of the other three inhibitors having a homophenylalanine at P1 site, *N*-Boc protected homophenyl alaninal was reacted with nitromethane and the nitroaldols were coupled with the corresponding peptide with free carboxylic terminus (Scheme 1, Table 1).

Scheme 1. Synthetic route for the preparation of the nitroalkene compounds used in this study.

103

104 **Inhibition of the M^{pro} activity by the peptidyl nitroalkenes.** The molecular structure and the K_i values

105 obtained for the inhibition of the M^{pro} activity by the six peptidyl nitroalkene compounds are summarized in

106 Table 1, respectively. For testing the inhibitory effect of the compounds enzymes from two different

107 expression systems were used (see Supporting Information). An example of the enzymatic activity inhibition

108 curves obtained using the enzyme obtained from the expression using the pMal-M^{pro} vector is shown in Figure

109 S1 in the Supporting Information. In Figure S2 are shown the inhibition profiles for the six compounds

110 obtained using the M^{pro} obtained from the expression using the pET21-M^{pro} vector. Compounds **FGA145**,

111 **FGA146** and **FGA147** with a glutamate surrogate at P1 site, leucine at P2 and an aromatic residue at P3

112 displayed inhibition at the low micromolar range, less than 10 μM. The *N*-terminal substitution of the

113 aromatic residue is well tolerated. The substitution of the benzyloxycarbonyl (Cbz) group of **FGA147** by a 4-

114 methoxy-1*H*-indole-2-carbonyl residue in **FGA146** leads to similar values of inhibition. Compounds **FGA86**

115 with homophenylalanine at P1 and phenylalanine at P2 sites, and an aromatic residue at P3; **FGA159** with a

116 homophenylalanine at P1 site, leucine at P2, and three more residues at P3, P4 and P5; and **FGA177** with a

117 homophenylalanine at P1 site, leucine at P2, and an aromatic residue at P3, also showed inhibition at the low

118 micromolar range. No irreversible character was observed over a 10-minute period (Figure S1 in the

119 Supporting Information) denoting the compounds to be non-time dependent inhibitors as it was predicted

120 by us.[24,25] The values of K_i obtained using the plate reader assay and the continuous fluorometric assay were

121 very similar, and the small non-significant differences might be due to differences in the enzymatic assays,

122 such as the use of different M^{pro} constructs and assay conditions, different amount of the organic solvent
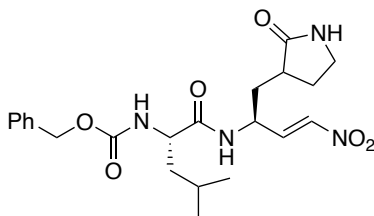
123 DMSO (Table 1).

Table 1 Structure and inhibitory activity against SARS-CoV-2 M^{pro} of the nitroalkene compounds.			
Compound	2D Structure	K_i (μM)[a]	K_i (μM)[b]
FGA86		17% inh. @ 20 μM	2.67 ± 0.29
FGA145		3.71 ± 0.38	9.82 ± 1.50
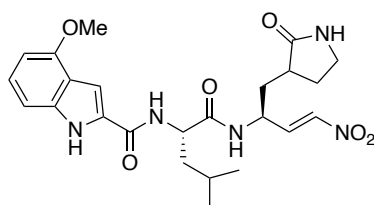
|			
FGA146		2.19 ± 0.18	0.96 ± 0.06
FGA147		2.18 ± 0.19	1.33 ± 0.05
FGA159		21.8 ± 1.5	1.19 ± 0.18
FGA177		25% inh. @ 20 µM	8.59 ± 1.60
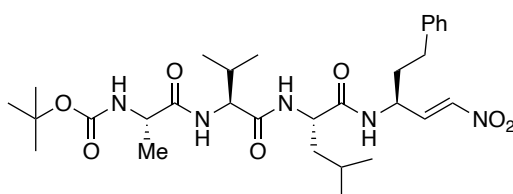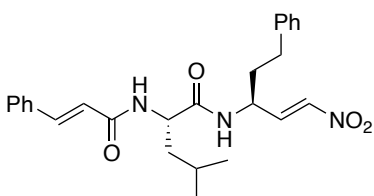
[a]Data obtained using the assay with M^{pro} obtained from the expression using the pMal-M^{pro} vector.
[b]Data obtained using the assays with M^{pro} obtained from the expression using the pET21-M^{pro} vector.

124

125 **Cellular antiviral activity and cytotoxicity.** Three compounds (**FGA145**, **FGA146** and **FGA147**) were selected

126 for the antiviral assay with infectious SARS-CoV-2. Huh-7-ACE2 cells were used for this antiviral assay. The

127 antiviral activity and cytotoxicity assays are shown in Figure 1. Compounds **FGA146** and **FGA147** showed

128 potent antiviral activity against with EC_{50} values in the low micromolar range (0.9 and 1.9 µM, respectively).

129 **FGA145** showed a less potent activity (EC_{50} = 11.7 µM), in line with the enzymatic activity inhibition results.

130 The cellular cytotoxicity of these compounds was very low. These three compounds were well tolerated with

131 CC_{50} values over 100 µM.

132

133
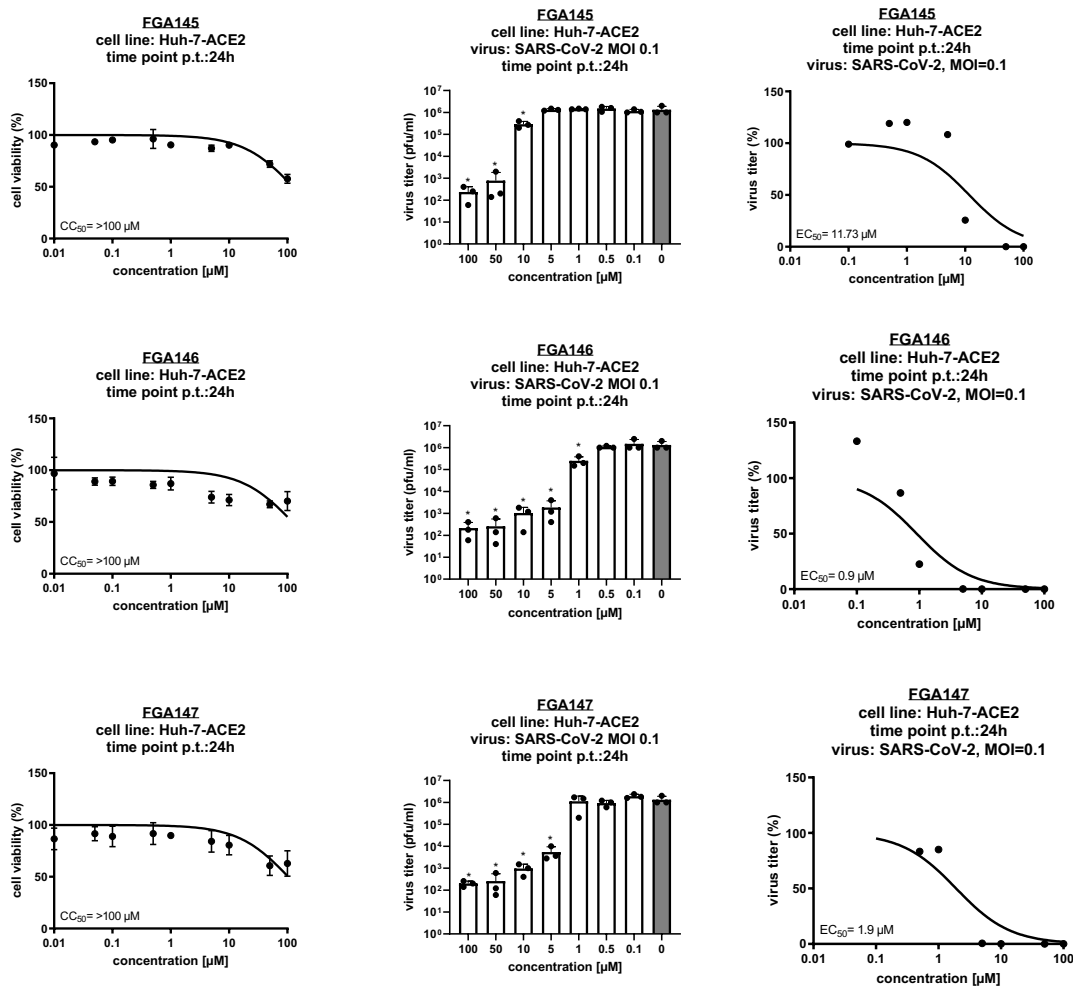
Figure 1. Inhibition of SARS-CoV-2 infection in Huh-7-ACE2 cell by FGA145, FGA146 and FGA147. Cytotoxicity assays for the three compounds (left column), all of them presented a CC_{50} greater than 100 μM. Effect of the three compounds on the virus titer (center and right columns); **FGA146** was the most potent inhibitor with a EC_{50} of 0.9 μM, followed by **FGA147** and **FGA145** with EC_{50} of 1.9 and 11.7 μM, respectively.

Inhibition of other proteases. While the cysteine protease M^{pro} is inhibited with high potency by inhibitors **FGA145**, **FGA146** and **FGA147**, no inhibitory activity against the serine proteases human matriptase (membrane-type serine protease 1, MT-SP1, prostamin) and bivalent expressed Zika Virus NS2B/NS3 (bZiPro) was observed by these compounds (Table 2). Besides inhibition of M^{pro}, compound **FGA145** was found to be a very potent inhibitor of cysteine proteases rhodesain (RhD), cruzain (CRZ), cathepsin L (CatL) and cathepsin B (CatB) with decreasing potencies from RhD to CatB (1.63 nM, 12.6 nM, 53.0 nM, 206 nM, respectively, Table 2). This finding is in line with previous reports of nitroalkenes as potent reversible inhibitors of these proteases.[23] Interestingly, all three compounds inhibit CatL, especially compound **FGA145**, denoting a possible multitarget effect of these compounds in the antiviral activity.

**Table 2 K_i values and selectivity towards some off-targets.**

Compound	MT-SP1	bZiPro	RhD (nM)	CRZ (nM)	CatL (nM)	CatB (nM)
FGA145	n.i.	n.i.	1.63 ± 0.22	12.6 ± 1.5	53.0 ± 4.1	206 ± 41
FGA146	n.i.	n.i.	n.d.	n.d.	868 ± 60	n.d.

7

FGA147	n.i.	n.i.	n.d.	n.d.	1993 ± 107	n.d.

n.i.: no inhibition (<10%) was observed at a concentration of 20 μM. n.d.: not determined. All data are mean values ± standard deviation of three technical replicates.
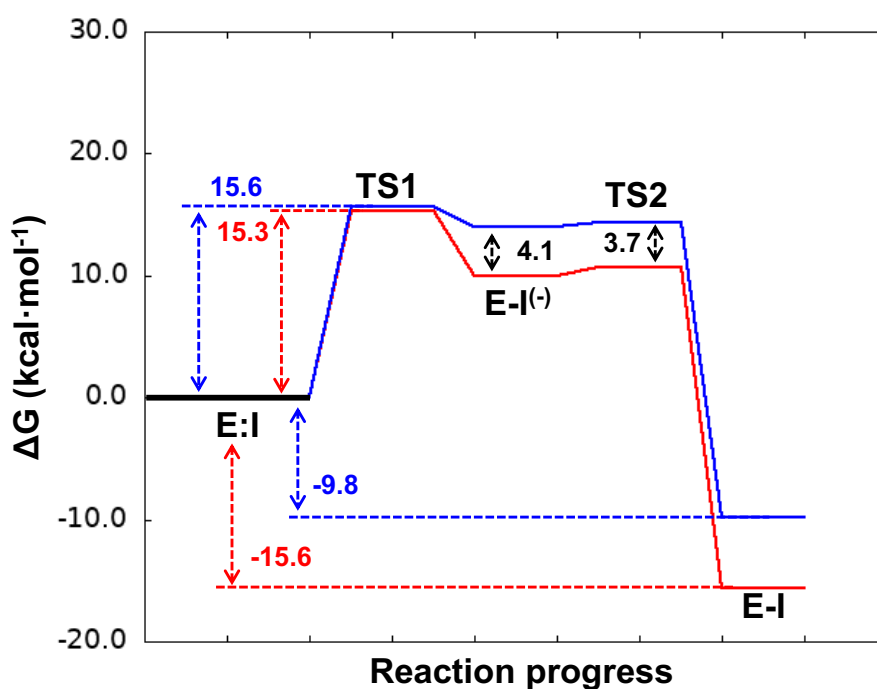
147

Computational study of the SARS-CoV-2 M^{pro} inhibition by FGA146 and FGA147. Based on the results derived from the kinetic studies, the inhibition reaction was studied according to the general mechanism proposed in Scheme 2 with the two most promising inhibitors: **FGA146** and **FGA147**. The first step would involve the activation of Cys145 by a proton transfer to His41 which take place concomitantly with the nucleophilic attack of the sulfur atom of Cys145 to the C20 atom of the inhibitor to form an intermediate, **E-I$^{(-)}$**. Then, the reaction is completed by the transfer of the proton from the protonated His41 to the Cα atom of the inhibitor to render the final **E-I** covalent adduct.
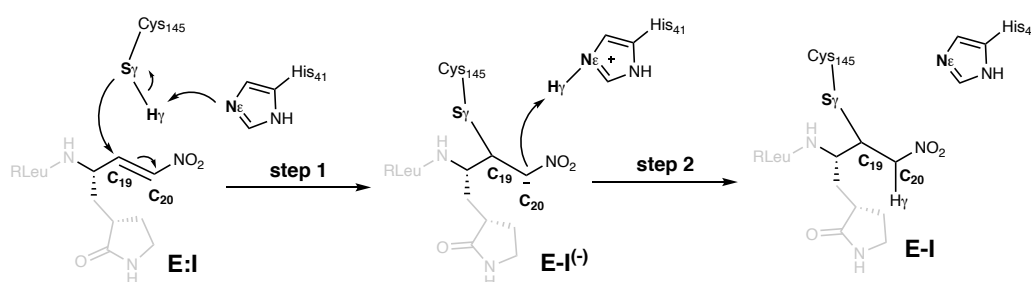
Scheme 2. Proposed Mechanism of SARS-CoV-2 M^{pro} Cysteine Protease Inhibition by nitroalkene compounds.

The free energy profiles depicted in Figure 2, derived from the free energy surfaces (FESs) of the SARS-CoV-2 M^{pro} inhibition with both **FGA146** and **FGA147** obtained by M06-2X/6-31+G(d,p)/MM MD simulations (see Figure S10 and S11 and computational details in the Supporting Information) confirm that the activation of Cys145 takes place concertedly with the inhibitor-enzyme covalent bond formation, **E:I** to **E-I$^{(-)}$** step. Interestingly, we already observed this concerted activation and nucleophilic attack of Cys145 when exploring the acylation step of the proteolysis reaction,[30] but previous inhibition processes explored in our laboratory have rendered stepwise processes where the activation of the Cys145 precedes the covalent formation between the sulfur atom of Cys145 and the different tested warheads of the inhibitors.[25,31] The second step of the inhibition reaction corresponds to the proton transfer from His41 to the C20 atom of the inhibitor, **E-I$^{(-)}$** to **E-I** step. As shown in Figure 2, the chemical step of the inhibition process is exergonic in both cases (-15.6 and -9.8 kcal·mol^{-1} with **FGA146** and **FGA147**, respectively), and the activation free energies, determined by the formation of the intermediate covalent intermediate, **E-I$^{(-)}$**, are very similar (15.3 and 15.6 kcal·mol^{-1} with **FGA146** and **FGA147**, respectively). These results are in agreement with the almost equivalent experimentally measured equilibrium K_i values (see Table 1), despite the simulations are done from **E:I** to **E-I**, and the measured K_i corresponds to the equilibrium from the solvated separated species **E + I**. The results suggest that the activity of the studied compounds is dictated by chemical steps of the full inhibitory process, with irrelevant effects of the binding step. A list of key inter-atomic distances obtained on the representative stable states is listed in Table S6.
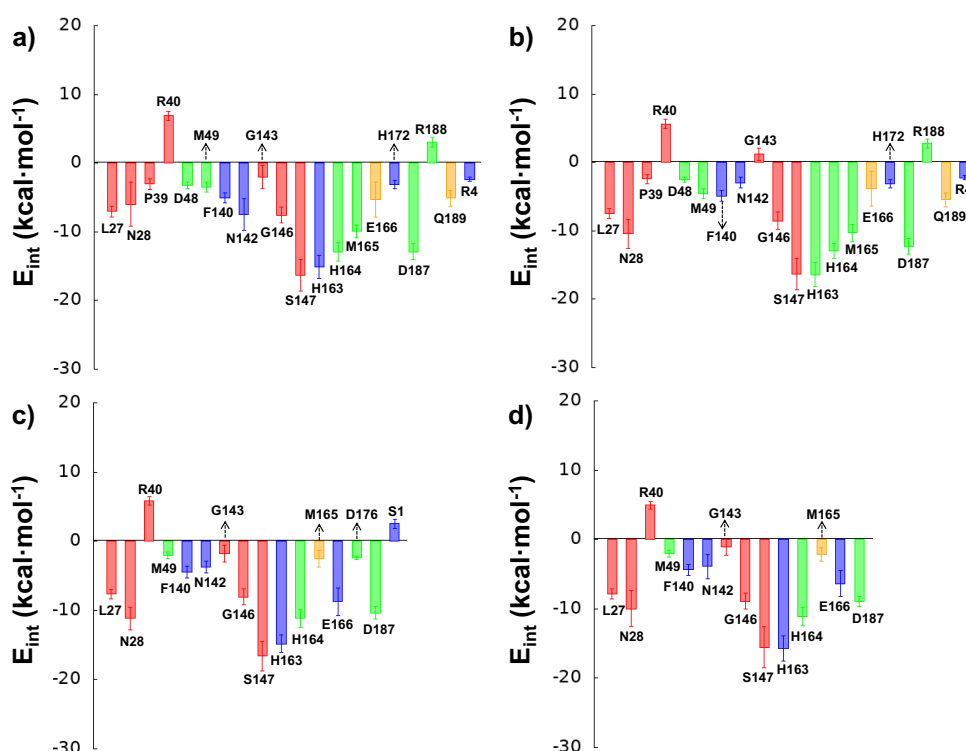
Figure 2. M06-2X/6-31+G(d,p)/MM free energy profiles for covalent complex formation between SARS-CoV-2 M^{pro} and: FGA146 (red line); and FGA147 (blue line) compounds. Energies are in kcal·mol^{-1}. The corresponding free energy surfaces are deposited in the Supporting Information.

The main interaction energies between residues of SARS-CoV-2 M^{pro} and the inhibitors **FGA146** and **FGA147**

computed in the **E:I** and in the **E-I** states are shown in Figure 3, while representative snapshots of the **E-I**

covalent product complexes are shown in Figure 4. It is important to point out that, due to the nature of the

computed interaction (electrostatic plus Lennard–Jones) some of the interactions do not necessary imply

close distance protein-inhibitor contacts. Consequently, these energetic results can complement structural

analysis. As observed in Figure 3, the pattern of interactions does not dramatically change from the Michaelis

complex **E:I** to the final **E-I** covalent product complex, in both reactions. Moreover, as reflected by the similar

plots obtained for **FGA146** and **FGA147**, the influence of the P3, the only fragment that distinguishes the two

inhibitors, does not dramatically affect the rest of the protein-inhibitor interactions, despite it has an effect

in the thermodynamics of the process, slightly more exergonic in the case of **FGA146** than in **FGA147**. The

interaction between the P3 moiety of **FGA146** with Glu166 and Gln189 residues are not observed in the case

of **FGA147**, where only a weak interaction has been measured with Met165 (see Figures 3 and 4). This is due

to the presence of the methoxy and the indole ring in the former that facilitates direct H-bond interactions

with polar residues of the cavity. These differences are related with the different orientation of the methoxy

indole ring of **FGA146** and the Cbz of **FGA147** in the S3 cavity, as discussed below. On the other side, in both

cases the interactions with S1' take place mainly through hydrogen bond interactions with the nitro oxygen

atoms of the P1 that is common in both inhibitors. However, while the nitro group of **FGA147** is interacting

with the oxyanion hole located in S1' formed by Gly143, Ser144, and Cys145, this interaction is partially lost

and an additional interaction with His41 is observed in the case of **FGA146** (see Figure 4). In addition, some

indirect interactions stabilize the P1' fragment, such as Leu27, Asn28, Pro39, Gly146, and Ser147, in **FGA146**, and Thr25, Asn28, Gly146, and Ser147 in **FGA147**. The specific favorable interactions between the lactam ring on P1 and S1 match in both inhibitors through interactions with F140, Asn142 and His163. The interaction with His172 is exclusive of the **FGA146** while the interaction with Glu166 is only observed in **FGA147**. His164 and Asp187 interact with P2 in both cases, while in the case of **FGA146** an additional interaction with Met165 is detected together with a weak interaction with Asp176. Finally, there are unfavorable interactions such as those between Arg40 and the warhead of both inhibitors, between Arg188 and P2 of **FGA146**, and between Ser1 of chain B and P1 in **FGA147**. Interestingly, in both complexes, Arg40 is ca. 9-10 Å from P1' while Arg188 is ca. 5-7 Å, thus corresponding to electrostatic interactions. Interestingly, the significant unfavorable interactions with Arg40 were already detected when studying the inhibition of M[pro] with other designed inhibitors.[31]

Figure 3. Main average interaction energies (electrostatic plus Lennard-Jones) between residues of SARS-CoV-2 M[pro] and the inhibitors FGA146 or FGA147, computed in the E:I and the E-I states. The colour of the bars indicate the specific interactions: red, blue, green and orange correspond to P1'∷∷S1', P1∷∷S1, P2∷∷S2 and P3∷∷S3 interactions, respectively. Results obtained as an average over 1000 structures of the AM1/MM MD simulations

216
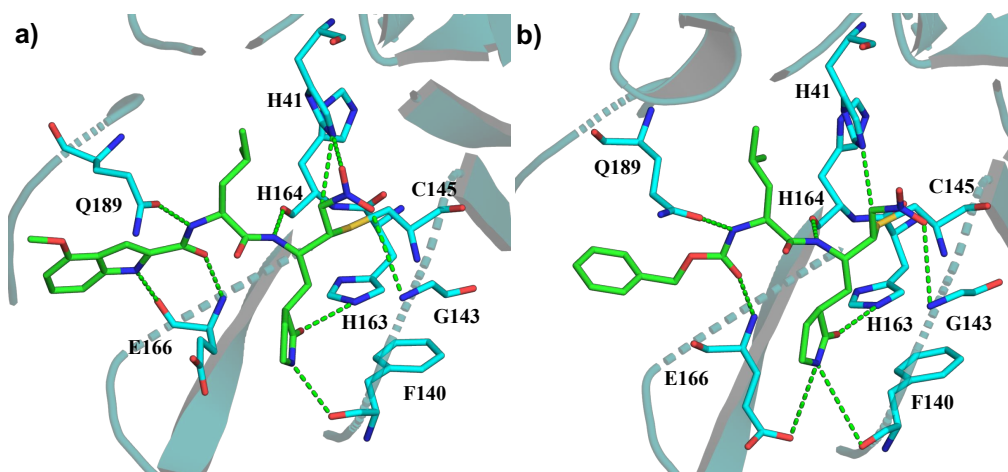
Figure 4. Snapshot of representative QM/MM MD structures of the E-I covalent product complex of SARS-CoV-2 M^{pro} with inhibitors (a) FGA146 and (b) FGA147. M^{pro} is shown in ribbon and the inhibitors in liquorice representation (in green). H atoms are omitted for clarity purposes.

220

Crystal structure of SARS-CoV-2 M^{pro} in the apo form and in complex with FGA146 and FGA147. The SARS-CoV-2 M^{pro} in complex with **FGA146** crystallized in the P2$_1$ space group and diffracted up to 1.98 Å resolution (Table S2) with one biological dimer in the asymmetric unit, and the complex with **FGA147** crystallized in the P2$_1$2$_1$2 space group and diffracted up to 1.62 Å resolution (Table S2) with one monomer per asymmetric unit that forms the biological dimer with a crystallographic-symmetry related neighboring molecule (Figure S4 in Supporting Information). The protein can be subdivided into three domains (as shown in Figure S4 in Supporting Information), domain I and domain II containing the active site and domain III is the dimerization domain.

After the structures of the respective protein complexes were solved, significant electron density was found at the active site. This electron density could be unequivocally assigned to the corresponding molecule inhibitors **FGA146** (Figure 5a) and **FGA147** (Figure 5b). Both inhibitors are covalently bound to the catalytic Cys145. A comparison of both inhibitors bound the active site is shown in Figure 5c.

233
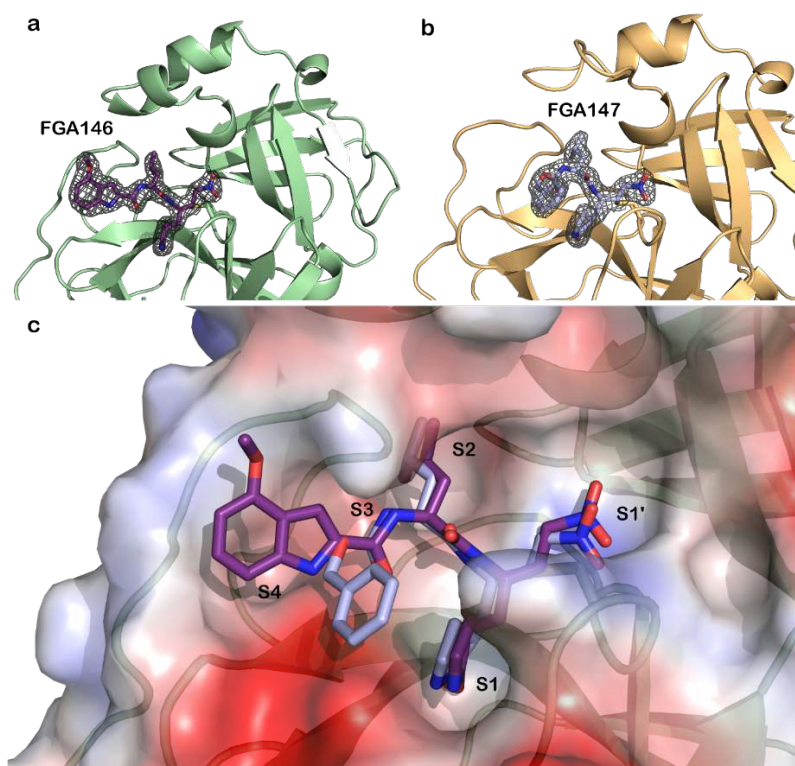
Figure 5. Crystal structures of SARS-CoV-2 M^pro in complex with inhibitors. M^pro is shown in ribbon and the inhibitors in ball and stick representation. 2Fo-Fc electron density map contoured at 1 (shown in gray mesh) of FGA146 (a) and FGA147 (b) bound covalently to the catalytic cysteine (Cys145). c Electrostatic surface representation of the active site of M^pro with bound FGA146 (violet) and FGA147 (light blue). Red indicates negative charge and blue positive charge.

The bound inhibitors, that mimic the natural peptide substrate, show a good geometric complementarity within the active site for subsites S1, S2 and S3 with the warhead located at the S1' subsite (Figure S5a,b,c in the Supporting Information). The nitro group of the warhead occupies, in both complexes, the "oxyanion hole" formed by the backbone amides of Gly143, Ser144, and Cys145 (Figure S5). The Sγ atom of the nucleophilic Cys145 forms a covalent bond to carbon C19 of the nitroalkene warhead of the inhibitor through a Michael addition (for numbering of the compounds see Figure S3). The S1 subsite is occupied by the glutamine surrogate γ-lactam ring that forms hydrogen bonds with the main chain of Phe140 and the side chain of Glu166 through the nitrogen atom (N16) of the ring, and to the side chain of His163 through the oxygen atom (O18) of the ring (Figure S5a,b,c and Figure S6b,d). The carbon atoms of the side chain of this residue lies in a hydrophobic cavity that forms the S1 subsite (Figure S5a,b,c and Figure S6b,d). The second residue of both inhibitors is a Leu that is inserted into the hydrophobic S2 subsite made up by His41, Met49, and Met169 (Figure S5a,b,c). There are three hydrogen bonds between the main chain of the peptidyl inhibitor and the protein. They involve interactions between atoms N10 from the inhibitors and O from His164, N3 and OE1 from Gln189, and O1 and N from Glu166. The P3 residue side chain of the inhibitors, a methoxy indole carbonyl group in **FGA146** and a Cbz group in **FGA147**, shows different conformations when bound to the protein (Figure 5c). The side chain of this residue from **FGA147** is oriented towards the solvent not having any interactions with protein residues (Figure 5c and S5a). The side chain of this residue from **FGA146** is occupying the S4 subsite (Figure 5c and S5a). This 4-methoxy-1H-indole-2-carbonyl group forms a

257 hydrogen bond between the N35 atom and the O atom from Glu166 (Figure S5b,d,e). This side chain is
258 encased inside the S4 subsite formed by residues Glu166, Leu167, Pro168, Gln189, Thr190 and Ala191 (Figure
259 S5d,e). The methoxy group is surrounded by the side and main chain of residue Gln189, and by the main
260 chain of residues Thr190 and Ala191 (Figure S5d,e). The distance between the O of this methoxy group and
261 the potential hydrogen bond partners is too far away and/or without a favorable geometry for this type of
262 interaction (Table S2), but it is enough to fix the position of this group and to orient the methyl group towards
263 the solvent. Interactions between protein atoms and inhibitor atoms and the distances between them are
264 summarized in Table S2. Besides the extensive hydrogen bond network, there are numerous non-polar
265 interactions that contribute to the tight binding of the inhibitor.

266 The position of the nitro group of the warhead in the structure of M^{pro} in complex with **FGA146** is not fixed
267 by the interactions with the residues that form the "oxyanion hole" (Gly143, Ser144 and Cys145, Figure S7).
268 In one of the monomers (mon. B), this nitro group forms hydrogen bond interactions with the N atoms from
269 Gly143 and Cys145 (Table S2, and Figure S7, protein in green and ligand in purple); while in the other
270 monomer (mon. A), the nitro group moves away from the "oxyanion hole" and interacts with His41 (Table S2,
271 and Figure S7, protein in cyan and ligand in brown); showing a certain degree of flexibility upon binding to
272 this S1 site.

273 Concerning the protein, there are different conformations of some residues forming the active site. The most
274 significant changes are located in the P2 helix, Ser46 to Asn51 (Figure 6a) and in the P5 loop, Asp187 to Ala193
275 (Figure 6a). The first segment is the α-helix that takes part in the formation of the S2 binding subsite. In the
276 structure of the **FGA146** complex, this segment is displaced towards the inside of the active site with respect
277 to the structure of the **FGA147** complex, consequently widening the S2 subsite (Figure 6a). The second
278 segment is the P5 loop, at this subsite only the inhibitor **FGA146** was observed to be bound, while the P3 side
279 chain of **FGA147** is pointing towards the solvent. This segment is tightly packed around **FGA146** fixing its
280 conformation. **FGA147** does not bind to this subsite in our structure, and the residues are positioned further
281 away (Figure 6a). Also, there is a small difference at the P4 β-hairpin flap. There is one loop close to the active
282 site with a different conformation (Figure 6a, loop I), and another loop further away from the active site also
283 with a different conformation (Figure 6a, loop II).

284 Besides the differences observed for the active site located in domains I and II, there are additional differences
285 in domain III between the structures of the M^{pro} in complex with **FGA146** and **FGA147** (Figure 6b). The
286 superposition of these structures shows clear displacements in the positions of four of the helices (α-6, α-7,
287 α-8 and α-9) in monomer A, as a consequence the loops linking those helices are also displaced.

288 These changes could not be observed in monomer B where all the helices from this domain are very well
289 aligned with only small non-significant differences (Figure 6b). Upon observing these differences in only one
290 monomer we superposed monomer A and monomer B from the M^{pro} in complex with **FGA146**, shown in
291 Figure 6c. Here, we were able to observe the same changes that occur in monomer A between the **FGA146**
292 and **FGA147** structures (Figure 6b). This indicates a great flexibility of the M^{pro} with no coordinated changes

293   in both monomers. These changes could be due to the different space groups in which the protein crystallized,

294   as stated previously [33] or just might indicate the flexibility of this protein which has to accommodate itself to

295   be able to catalyze the proteolysis of the polyprotein to liberate the mature proteins essential for the virus

296   replication.

297

298
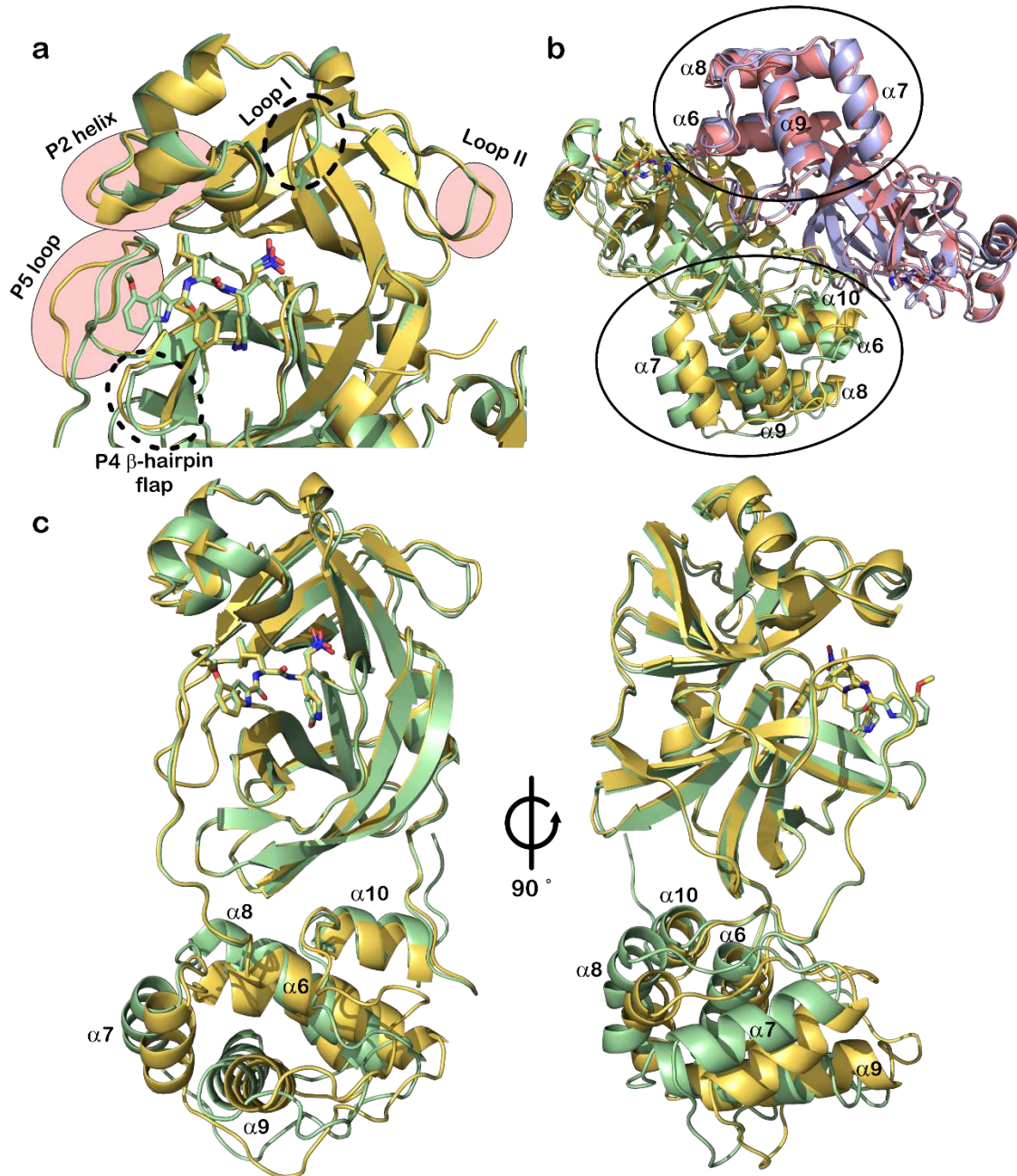
Figure 6. Conformational changes in M^{pro} upon binding of the inhibitors. a Most significant changes in the active site and domain I are located at P2 helix and the P5 loop. Smaller changes can be observed at the P4 β-hairpin flap, loop I and loop II. b Superposition of M^{pro} in complex with FGA146 and FGA147. Monomer A and B from the complex with FGA146 are shown in light green and yellow orange, respectively, and from the complex with FGA147 are shown in salmon and light blue, respectively. Significant displacements of some of the helices from the dimerization domain (Domain III, circles) of monomer A can be observed, while the same domain from monomer B does not show these displacements. c Superposition of monomer A (light green) and B (yellow orange) of the M^{pro} in complex with FGA146. The helical dimerization domain (Domain III) shows differences in the relative positions of some of the helices. Four

14

307 **of them show significant displacements (α-6, α-7, α-8 and α-9) while the las helix (α-10) does not show any significant**
308 **displacement.**

309

310 **Comparison of computational and crystallographic structures.** The X-ray structures confirm the predictions

311 based on the QM/MM simulations. As shown in Figure 7 and 8, the comparison of structures derived from

312 computer simulations with the corresponding complexes derived from experiments confirm a good

313 agreement, both techniques virtually describing the same binding mode of the most active compounds in the

314 active site of M^{pro}. A detailed analysis of the active site of the SARS-CoV-2 M^{pro} in complex with inhibitors

315 **FGA146** and **FGA147** can be derived from Figures 3-8 and Figures S5-S6, together with interatomic distances

316 between protein atoms and inhibitor atoms that are summarized in Table S2 and S6. It is important to stress

317 that the computer simulations were initiated from a previous solved X-ray structure of SARS-CoV-2 M^{pro}

318 complexed with the N3 inhibitor (PDB ID 6LU7).[22]
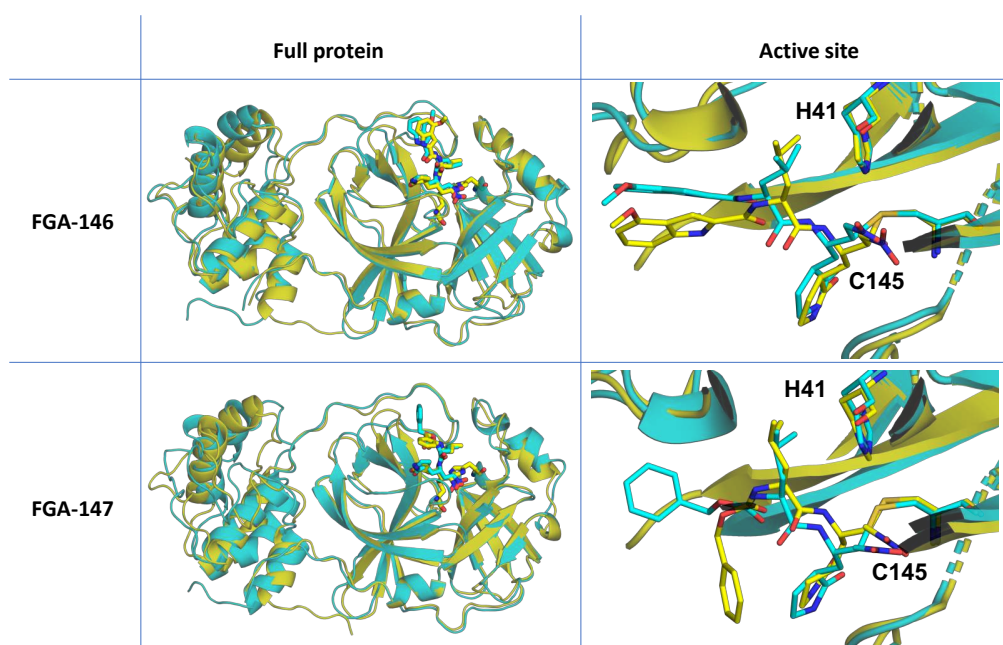
319

320 **Figure 7. Detail of the FGA146 and FGA147 inhibitors covalently bounded to the active site of SARS-CoV-2 M^{pro} through**
321 **Cys145. Results derived from X-ray diffraction (in yellow) and QM/MM MD studies (in cyan).**

322

323

15
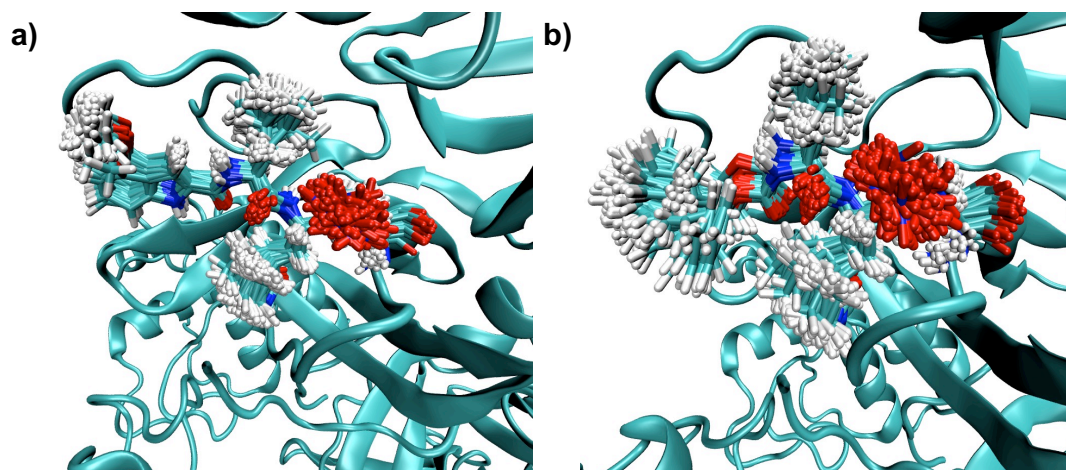
Figure 8. Overlap of structures generated during the MD simulation of the (a) FGA146 and (b) FGA147 inhibitors, covalently bounded to the active site of SARS-CoV-2 M^{pro} through Cys145. Structures randomly selected every 100 frames for clarity purposes.

The computational findings of **E-I** product complex of the inhibition with **FGA146** fit with the presence of two different conformations in monomer A and B in the crystal structure. Thus, the nitro group can slightly move away from the "oxyanion hole" and approaching to His41 during the MD simulations, mirroring the crystallographic data of monomer A (Table S6 and Figures 3a and 8a vs Table S2 and Fig. S7, cyan). In the case of **FGA147**, the nitro group appears to interact basically with the oxyanion hole in both, simulations and X-ray diffractions. Regarding the **FGA147**, an *a priori* discrepancy between experiments and theory is found regarding the P3 residue side chain of the Cbz group. Thus, the QM/MM MD simulations suggest the P3 residue is oriented, on average, towards the S4 subsite (Figure 4b and 7). This most populated orientation of P3 in **FGA147** is stabilized by an interaction between the phenyl group of Cbz and the Gln189 and Glu166 residues, which is confirmed by the computed favorable interaction (Figure 3). On the contrary, the X-ray structure of M^{pro} complex with **FGA147** shows that side chain of this residue is oriented towards the solvent not having any interactions with protein residues (Figures 5 and 7). In this case, the position of the Cbz group in **FGA147** is forced by a symmetry-related molecule into the crystal, located just above the active site and thus preventing the inhibitor to adopt a conformation that matches the species suggested by computer simulations. However, when considering the conformational space explored during the MD simulations (Figure 8b), orientations of P3 similar to the one determined by means of X-ray diffraction were detected (Figure 8b). In the case of **FGA146**, the orientations observed for the 4-methoxy-1H-indole-2-carbonyl group appear to be more constrained (Figure 8a). Interestingly, the two possible orientations of P3 in **FGA147** have been detected in previously solved structures of SARS CoV-2 M^{pro} in complex with peptidomimetic covalent inhibitors presenting the same recognition part as **FGA147** (see Figure S12).[34-38] Concerning the rest of the protein (Figure 7), there is a general good agreement between experiments and simulations, except for the very flexible regions exposed to the solvent (the C-terminal loop, S301-G302-V303-T304-F305-Q306 residues). These results confirm the robustness of the results and the low impact of the ligand on the full structure of the protein.

Thermal stability of M^{pro} in the absence and presence of the inhibitors. We have examined the thermal stability of the protease in the absence and in the presence of the inhibitors using circular dichroism (CD). The T_m value for M^{pro} did not change, significantly, in the presence of **FGA177** (0.1 ± 0.2 °C in the range of 25 to 100 μM, Figure 9c and S8d). In the presence of **FGA86** (Figures 9c and S8a) at the lower concentration (25 μM) there was a slight increase in stability (0.9 ± 0.4 °C) and at the higher concentration (100 μM) there was a slight decrease in stability (-1.7 ± 0.4 °C). The presence of the inhibitors **FGA145** (-1.1 ± 0.2 °C at 25 μM and -3.7 ± 0.1 °C at 100 μM, Figures 9c and S8b) and, while **FGA146** (-4.2 ± 0.2 °C at 25 μM and -6.4 ± 0.1 °C at

16

360   100 μM, Figure 9a,c) and **FGA147** (-3.0 ± 0.1 °C at 25 μM and -4.2 ± 0.2 °C at 100 μM, Figure 9b,c) lead to a

361   significant decrease of the T_m value. It has been reported that the association of covalently-bound compounds

362   induce shifts of T_m values to lower temperatures with an apparent destabilization of the protein.[39-41] Thus,

363   these data agree with the formation of a covalent bond between the protein and compounds **FGA86, FGA145**,

364   **FGA146** and **FGA147**. Based on the X-ray crystallographic analyses and thermal stability data, we can conclude

365   that these last four compounds bind covalently to the protein and are potent inhibitors of its enzymatic

366   activity. The inhibitor **FGA159** showed an increase in stability of 2.5 ± 0.2 °C at a concentration range between

367   5 and 100 μM (Figures 9c and S8c), this might indicate the formation of reversible covalent interactions.[42]

368
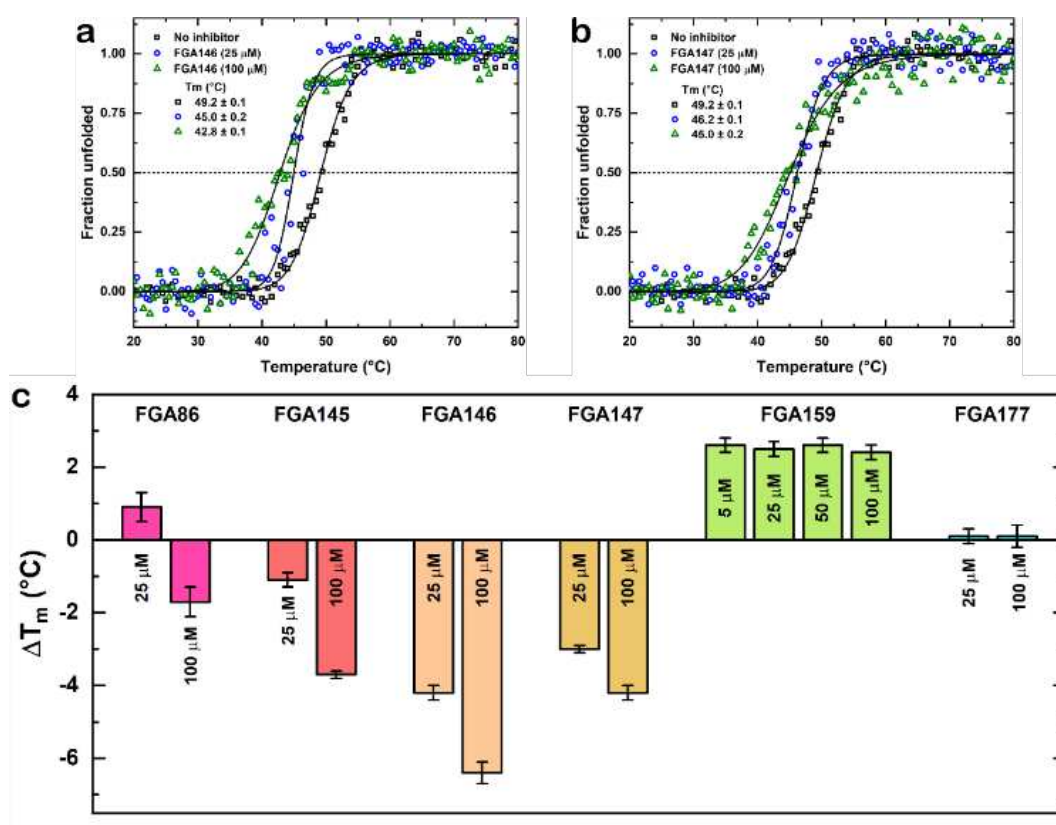
369

370   **Figure 9. Effect of the inhibitors on the thermal stability of M^{pro}. a Thermal stability of M^{pro} in the presence of FGA146**
371   **using circular dichroism. b Thermal stability of M^{pro} in the presence of FGA147 using circular dichroism. The T_m value**
372   **of M^{pro} in the absence of inhibitors (black squares) was 49.2 °C, while in the presence of 25 (blue circles) and 100 μM**
373   **(green triangles) of FGA146, the values decreased to 45.0 and 42.8 °C, respectively. In the case of FGA147, these values**
374   **decreased to 46.2 and 45.0 °C, respectively. c Change of the T_m in the presence of the inhibitors. FGA86 shows a slight**
375   **increase in stability at the lower concentration (25 μM) and a decrease in stability at the higher concentration (100**
376   **μM). There is no significant change in stability in the presence of FGA177; while in the presence of FGA145, FGA146**
377   **and FGA147 shows a significant decrease in stability, suggesting that they could bind covalently to the protein. The**
378   **presence of FGA159 increased the stability of the protein, indicating the possible formation of reversible non-covalent**
379   **interactions.**

380

381   **Binding of inhibitors to M^{pro}.** Figure 10a shows the binding isotherms of **FGA145** and **FGA146** to M^{pro} by

382   ultracentrifugation, and Figure 10b shows the binding of **FGA147** to M^{pro} measured by ITC. The measured

383   binding dissociation constants for these compounds were in the low micromolar range, the K_d for **FGA145**

384   was 11.8 ± 1.05 μM, the K_d for **FGA146** was 7.28 ± 0.58 μM (the average obtained from using both, absorption

17

385   and fluorescence, data), and the K_d for **FGA147** was 2.86 ± 0.25 µM. For all three inhibitors we obtained a

386   stoichiometry of one, which is compatible with binding of one molecule of the inhibitor to the active site.

387

388
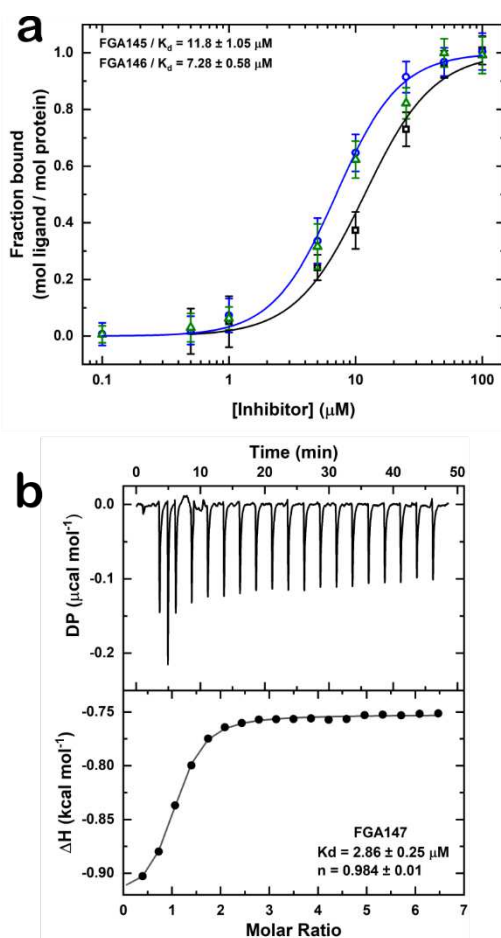
Figure 10. Binding of inhibitors to M^{pro}. a Binding isotherm of FGA145 (black squares) and FGA146 (blue circles and green triangles) to M^{pro}. The concentration of FGA145 was measured by absorption, and that of FGA146 was measured by absorption (blue circles) and fluorescence (green triangles). b ITC binding profile of FGA147 to M^{pro}.

392

Discussion

394   We have designed, synthesized and measured the inhibitory effect of a series of peptidomimetic compounds

395   with a nitroalkene warhead on the enzymatic activity of M^{pro} and cell infection. We have also examined the

396   possibility of using a nitroalkene warhead that due to its reversible binding [23] should decrease the possibility

397   of side effects due to unwanted reactions with other cellular components.

398   Six compounds (**FGA86**, **FGA145**, **FGA146**, **FGA147**, **FGA159** and **FGA177**) were prepared in good yields via

399   a short and straightforward synthetic route. All of them exhibited enzyme inhibitory activity (K_i: 1-10 µM) and

400   three of them (**FGA145**, **FGA146** and **FGA147**), having the typical coronaviral protease glutamine surrogate

401   (beta-lactam) at P1 site and a L-leucine at P2 site, gave good anti-SARS-CoV-2 infection activity in the low

402   micromolar range (EC$_{50}$: 1-12 µM) without significant toxicity. Additional kinetic studies of the selectivity of

403   **FGA145**, **FGA146** and **FGA147** show that they are also potent inhibitors of cathepsin L (CatL), revealing a

404   multitarget effect.

18

405 QM/MM computer simulations assisted in elucidating the way of action of the most promising compounds,

406 **FGA146** and **FGA147**, by generating the complete free energy landscape of the inhibitor-enzyme covalent

407 complex formation. The results of the inhibitory mechanism, that appear to be equivalent in both cases,

408 suggest that activation of Cys145 takes place concertedly with the inhibitor-enzyme covalent bond formation.

409 In the second step, the final proton transfer takes place from His41 to the Cα atom of the inhibitors, E-I$^{(-)}$ to

410 E-I step. The resulting free energy profiles for the covalent inhibition of SARS-CoV-2 M^{pro} with **FGA146** and

411 **FGA147** show how the processes are exergonic in both tested inhibitors, determined by the first step. This

412 indistinguishable predicted kinetic and thermodynamic behavior of **FGA146** and **FGA147** agrees with the very

413 close experimentally determined K_i values (see Table 1), thus suggesting that the inhibitory activity of the

414 inhibitors can be dictated by the chemical steps of the inhibition process. QM-MM averaged interaction

415 energies (electrostatic plus Lennard–Jones) between residues of SARS-CoV-2 M^{pro} and the inhibitors **FGA146**

416 and **FGA147** allows complementing the geometrical analysis based just on short distance (H-bond)

417 interactions. The similar plots obtained for **FGA146** and **FGA147**, indicate that the influence of the P3, that is

418 the only fragment that differentiate them, does not dramatically affect the rest of the protein-inhibitor

419 interactions, despite having a significant effect in the reactivity.

420 Finally, the crystal structures of the M^{pro} in complex with **FGA146** and **FGA147** were solved and confirmed the

421 binding modes. These binding modes agree with the computer predictions. The covalent inhibitory character

422 of these inhibitors is similar to other peptidomimetic inhibitors.[18, 19, 22] Our crystal structures, that virtually

423 overlap with the structures derived from the computer simulations that were initiated from a previously

424 crystallized M^{pro} in complex with a different inhibitor (N3), corroborates the great conformational flexibility

425 of the dimer of M^{pro}. Flexibility of the active site has been reported to be needed to accommodate the

426 different natural cleavage sites present in the polyprotein of SARS-Cov-2, as well as some conformational

427 flexibility of other regions of the protein (Domain II and III).[33]

428 Through thermal denaturation we have been able to observe that the inhibitors might stabilize or destabilize

429 the protein, in some cases a destabilization of more than 6 °C was found. The analysis of the crystal structures

430 together with these thermal denaturation data shows the influence of the inhibitors on the whole structure

431 of M^{pro}. This instability induced by active site inhibitors might be exploited to increase their potency against

432 the virus replication if they could be combined with inhibitors that bind to sites other than the active site to

433 further disrupt the activity of this essential protease for the virus.

434 In summary, we have designed and synthesized six compounds as inhibitors of the SARS-Cov-2 M^{pro} with the

435 nitroalkene warhead. The three most active inhibitors were active at low micromolar concentrations against

436 virus and did not show significant toxicity. These compounds were also active against human Cathepsin L in

437 the nanomolar range, denoting a dual activity. The fact that they are reversible covalent inhibitors would

438 decrease the possibility of side effects due to unwanted reactions with other cellular components. The

439 computer simulations and the crystal structures of the two most promising inhibitors in complex with M^{pro}

440 show the mechanism of action of these inhibitors and the interactions established between the inhibitor and

441   the protein. All these results combined suggest the viability of employing these compounds as promising

442   drugs against SARS-CoV-2 and new coronavirus that might appear in the future.

443

Methods

444 **Methods**

445

446 **General procedure for the preparation of nitroalkenes**

447   To a stirred solution of alcohol **1** (1.05 g, 4 mmol) in dichloromethane (32 mL) was added Dess-Martin

448   periodinane (1.82 g, 4.3 mmol) and sodium bicarbonate (361mg, 4.3 mmol). The resulting mixture was stirred

449   at room temperature for 1 h. Then the reaction mixture was cold with an ice-bath and triethylamine (0.17

450   mL, 1.21 mmol) and nitromethane (1.33 mL, 24.4 mmol) were. Then the mixture was stirred for 15 h at room

451   temperature and then was quenched with a saturated aqueous solution of NH_4Cl (10 mL), the mixture was

452   extracted with CH_2Cl_2 (3 x 15 mL) and the combined organic layers were washed with HCl 1M, then with a

453   saturated aqueous solution of sodium bicarbonate and then dried over Na_2SO_4. Then the solvent was

454   evaporated and the residue was purified by column chromatography (silica gel, CH_2Cl_2/MeOH (99:1 to 85:15)

455   to afford the desired product as a yellow oil (71%).

456   The corresponding nitroaldol (0.73 mmol) was dissolved in dichloromethane (2.1 mL) and placed in an ice

457   bath. Then trifluoroacetic acid (1.1 mL) in dichloromethane (1.1 mL) was added dropwise and the mixture

458   was stirred at room temperature for 3 h. The reaction mixture was evaporated in vacuo to give the product

459   as a colorless solid. The resulting mixture was submitted to the next step without any further purification.

460   To a solution of the ammonium trifluoroacetate (0.80 mmol) and the carboxylic acid (0.89 mmol) in

461   dichloromethane (8 mL) cold with an ice-bath, HOBt·$H_2$0 (121 mg, 0.89 mmol) was added. After 15 min at the

462   same temperature, DIPEA (0.56 mL, 3.23 mmol) was added dropwise. After another 15 min, EDC (186.2 mg,

463   0.97 mmol) was added and the mixture was stirred for 16 h at room temperature. Then the mixture was

464   quenched with saturated ammonium chloride solution (10 mL) and extracted with dichloromethane (3 x 20

465   mL). The combined organic layers were washed with HCl 1M, with a saturated aqueous sodium bicarbonate

466   solution and then dried over Na_2SO_4. Then the solvent was evaporated and the residue was purified by

467   column chromatography (silica gel, CH_2Cl_2/MeOH (100:0 to 85:15) to afford the desired product (64%, two

468   steps).

469   To an ice bath cold solution of peptidyl nitroaldol (0.66 mmol) in dichloromethane (6.6 mL) was added DIPEA

470   (0.24 mL, 1.39 mmol), then methanesulfonyl chloride (0.056 mL, 0.73 mmol). The resulting mixture was

471   stirred overnight, then it was quenched with a saturated aqueous solution of NH_4Cl (10 mL) and extracted

472   with dichloromethane (3 x 20 mL). The combined organic layers were washed with HCl 1M then with a

473   saturated aqueous sodium bicarbonate solution and then dried over Na_2SO_4. Then the solvent was

474   evaporated and the residue was purified by column chromatography (silica gel, CH_2Cl_2/MeOH (99:1 to 9:1)

475   to afford the desired product (68-81%).

476   For the preparation of all the compounds, the coupling steps and nitroalkene formation were done following

477   the experimental procedure detailed above.

478   For the preparation of compound **FGA159**, the hydrolysis and hydrogenation steps were done following

479   standard experimental procedures.

480

481 **Cloning of M^{pro} gene.** The M^{pro} gene was cloned in two different vectors with a similar strategy. First, the

482   sequence of the gene coding for M^{pro} (nsp5) SARS-CoV-2 was optimized for *Escherichia coli* expression and

483   synthesized and cloned directly into pUCIDTKan vector by the company Integrated DNA Technologies (IDT,

484   Leuven, Belgium) and named pUCIDTKan-M^{pro}. The M^{pro} gene was amplified from the vector pUCIDTKan-M^{pro}

485   using the following primers: 5'-CGGGCGC<u>CATATG</u>TCTGCTGTTCTGCAGAGTG-3' (NdeI site) and 5'-

486   CCG<u>CTCGAG</u>TTAATGGTGATGGTGATGG-3' (XhoI site) and cloned into the vector pET21a (Novagen) named

487   pET21-M^{pro}. The cloned gene possess one M^{pro} autocleavage site SAVLQ↓SGFRK (arrow indicates the

488   cleavage site) at the N-terminus, and at the C-terminus, the construct codes for the human rhinovirus 3C

489   PreScission protease cleavage site (SGVTFQ↓GP) connected to a His6 tag. The authentic N-terminus is

490   generated by M^{pro} autoprocessing during expression, whereas the authentic C-terminus is generated by the

491   treatment with PreScission protease. Second, the M^{pro} gene was inserted into the pMal plasmid harboring

492   the C-terminal hexahistidine-tagged sequence of SARS-CoV-2 Mpro named pMal-M^{pro} (Prof. John Ziebuhr,

493   Justus Liebig University Gießen, Germany). The sequence contained the native nsp4/nsp5 M[pro] cleavage site
494   between MBP and Mpro as well as the native nsp5/nsp6 cleavage site between M[pro] and the hexahistidine
495   tag, thus enabling the purification of native Mpro.
496
497   **Protein expression and purification.** *SARS-CoV-2 M[pro].* The vector pET21-M[pro] was transformed into *E. coli*
498   Tuner (DE3) cells (Novagen, Merck, Madrid, Spain). These cells were grown in 2xYT medium supplemented
499   with ampicillin (100 mg/L) at 37 °C. When the cells attained an OD_{600} of 0.6—0.8 the temperature was then
500   dropped to 20 °C. When the temperature stabilized (approx., 15 min) the expression of the protein was
501   induced by the addition of 0.1 mM isopropyl-β-D-thiogalactopyranoside (IPTG), and let to grow for an
502   additional 16 h. The cells were harvested by centrifugation and resuspended in the lysis buffer containing 10
503   mM TRIS-HCl at pH 8.0 and 1% (v/v) TritonX-100. The cells were lysed by sonication and the insoluble fraction
504   was removed by centrifugation at 45,000 × g for 1 hour; the supernatant was then loaded onto a HisTrap FF
505   column (GE Healthcare). The column was washed extensively first with 10 mM TRIS-HCl buffer at pH 8.0
506   containing 5 mM imidazole, 0.5 M NaCl and 2 mM β-mercaptoethanol (β-ME) and subsequently with 10 mM
507   TRIS-HCl buffer at pH 8.0 containing 50 mM imidazole, 0.5 M NaCl and 2 mM β-ME. The protein was eluted
508   from the column in 10 mM TRIS-HCl buffer at pH 8 containing 300 mM imidazole, 0.5 M NaCl and 2 mM β-
509   ME. Just after the elution of the protein, the concentration of the β-ME was raised to 10 mM. The fractions
510   containing the protease were then pooled, and PreScission protease containing a hexahistidine tag was
511   added at a 500:1 molar ratio. The mixture was then dialyzed against a solution containing 10 mM TRIS-HCl at
512   pH 8, 100 mM NaCl and 2 mM β-ME for 18 h at 4 °C. The PreScission-treated M[pro] solution was applied to a
513   HisTrap FF column to remove the PreScission protease, the C-terminal tag, and M[pro] with uncleaved
514   hexahistidine tag. The processed M[pro] was collected in the flow-through and concentrated to 10 mg/mL. The
515   generation of the proper N-terminal residue was confirmed by N-terminal sequencing by Edman degradation
516   performed by the Protein Chemistry facility of the Centro de Investigaciones Biológicas
517   (https://www.cib.csic.es/facilities/scientific-facilities/protein-chemistry). The expression of SARS-CoV-2 M[pro]
518   using the vector pMal-M[pro] was performed exactly as described previously[43]. *Human matriptase.*
519   Recombinant expression and purification was mainly performed as described previously[44]. The pQE30
520   plasmid, containing the human matriptase (membrane-type serine protease 1, MT-SP1, prostamin) was
521   kindly provided by Prof. Torsten Steinmetzer (Philipps University Marburg, Germany). Since MT-SP1 is
522   expressed as inclusion bodies, no leakage suppression was needed and, hence, the plasmid was transformed
523   in *E. coli* BL21-Gold (DE3) (Agilent Technologies, Santa Clara, CA, USA) cells. After growing them in LB medium
524   supplemented with ampicillin (100 mg/mL) to an OD_{600} of 0.6—0.8, overexpression was induced by addition
525   of 1 mM IPTG over night (o.n.) at 20 °C. Cells were harvested by centrifugation, flash frozen in liquid N_2 and
526   stored at −80 °C until further usage. For protein refolding and purification from inclusion bodies, cell pellets
527   were resuspended in lysis buffer (50 mM TRIS−HCl pH 8.0, 300 mM NaCl, 10% (v/v) glycerol and 1 mM β-ME),
528   supplemented with lysozyme and DNase and stirred for 1 h at room temperature (rt). After that, cells were
529   further lysed by sonication (Sonoplus HD 2200; Bandelin, Berlin, Germany) and again centrifuged. The
530   supernatant was discarded and the pellet was washed with lysis buffer. Proteins were solubilized in a
531   denaturing solubilization buffer (50 mM TRIS−HCl at pH 8.0, 6 M urea, 10% (v/v) glycerol and 1 mM β-ME) by
532   stirring o.n. at rt. The suspension was again centrifuged to remove cell debris. The supernatant was subjected
533   to IMAC on a HisTrap HP 5 ml column (Cytiva Europe GmbH, Freiburg im Breisgau. Germany), using IMAC
534   buffer A (50 mM TRIS−HCl pH 8.0, 6 M urea, 20 mM imidazole and 1 mM β-ME) in a linear gradient with IMAC
535   buffer B (50 mM TRIS−HCl at pH 8.0, 6 M urea, 250 mM imidazole and 1 mM β-ME). The fractions, containing
536   eluted MT-SP1 were refolded by a 2-step dialysis over 12 h each at 4 °C in dialysis buffer A (50 mM TRIS−HCl
537   at pH 9.0, 3 M urea and 1 mM β-ME) and anion exchange (IEX) buffer A (50 mM TRIS−HCl at pH 9.0, 1 mM β-
538   ME) prior to IEX chromatography on a Resource Q 1 ml column (Cytiva Europe GmbH, Freiburg im Breisgau.
539   Germany), using IEX buffer A in a linear gradient with IEX buffer B (50 mM TRIS−HCl at pH 9.0, 1 M NaCl and
540   1 mM β-ME). Eluted MT-SP1 was flash frozen in liquid N_2 and stored at −80 °C. *Zika Virus 2 NS2B[CF]/NS3[pro].*
541   The bivalently expressed ZIKV protease was expressed and purified as described previously[45]. Briefly, the

21

542 pETDUET vector containing bZiPro (purchased from Addgene) was transformed into competent *E. coli* BL21
543 Gold (DE3) cells (Agilent Technologies, Santa Clara, CA, USA) and grown in LB medium containing ampicillin
544 at 37 °C until they attained an optical density (OD_{600}) of 0.8. Overexpression was induced o.n. by addition of
545 1 mM IPTG at 20 °C. After harvesting, cells were flash frozen in liquid N_2 and stored at −80 °C until protein
546 purification. Herein, cell pellets were resuspended in lysis buffer (20 mM TRIS−HCl at pH 8.0, 300 mM NaCl,
547 20 mM imidazole, 0.1% (v/v) Triton X-100, RNase, DNase, lysozyme and 1 mM DTT) and lysed by sonication.
548 After centrifugation, bZiPro from the cleared supernatant was purified by IMAC on a HisTrap HP 5 ml column
549 with a step-gradient of washing buffer (20 mM TRIS−HCl at pH 8.0, 300 mM NaCl and 20 mM imidazole) and
550 elution buffer (20 mM TRIS−HCl at pH 8.0, 300 mM NaCl and 250 mM imidazole). The eluted fractions,
551 containing bZiPro were subjected to a gel filtration step (HiLoad 16/600 Superdex 75; GE Healthcare, Chicago,
552 IL, USA) in SEC buffer (50 mM TRIS−HCl at pH 8.0 and 150 mM NaCl). Eluted bZiPro was flash frozen in liquid
553 N_2 and stored at −80 °C. ***Cruzain.*** Cruzain (CRZ) was kindly provided by Dr. Avninder S. Bhambra (De Montfort
554 University, Leicester, UK). ***Cathepsin L, Cathepsin B.*** Both cathepsin L (CatL) and cathepsin B (CatB) were
555 purchased from Calbiochem (Merck Millipore, Burlington, Massachusetts). ***Rhodesain.*** Rhodesain (RhD) was
556 recombinantly expressed and purified as reported previously[46].

557

558 **Enzymatic assays.** Proteolytic activity was determined by cleavage of fluorescence resonance energy transfer
559 (FRET) peptide substrates. Fluorescence was measured using a Fluorolog-3 (Horiba Jobin Yvon, France)
560 photon counting spectrofluorometer and a TECAN Infinite F2000 PRO plate reader (Agilent Technologies,
561 Santa Clara, USA). The Fluorolog-3 spectrofluorometer was used to measure the activity of M^{pro} using the
562 substrate Dabcyl-KTSAVLQ↓SGFRKME-(Edans)-Amid (Biosyntan, Berlin, Germany). They were carried out
563 with excitation wavelength of 360 nm (4 nm band pass) and emission wavelength of 460 nm (8 nm band pass)
564 using 5 x 10 mm cells at 25 °C, a protein concentration of 0.05 mg/mL (1.47 μM) and a substrate
565 concentration of 5 μM with various concentrations of the inhibitors. The inhibitor was added into the M^{pro}
566 solution in the reaction buffer, mixing and allowing the mixture to equilibrate for 10 sec, and then initiated
567 by adding the substrate solution. Compounds in Table 1 were diluted in DMSO (**FGA145, FGA146** and
568 **FGA147**), N,N-dimethylformamide (**FGA86** and **FGA177**) and ethanol (**FGA159**). Due to the deleterious effect
569 of the solvents on the activity of M^{pro} and for consistency of the data, the concentration of solvent was kept
570 constant at 1% (v/v) in all experiments. The fluorescence time course of the reaction mixture was recorded
571 continuously for 2 min in 10 mM TRIS-HCl buffer at pH 8.0 containing 0.1 M NaCl and 2 mM β-ME. The slope
572 of the curve of fluorescence intensity with time quantitatively reflects the activity of the enzyme. The
573 proteolytic reaction initial velocity in the presence or absence of the inhibitors was determined by linear
574 regression using the data points from the first 10 sec of the kinetic progress curves. The IC_{50} was calculated
575 by adjusting a sigmoidal curve to the initial velocities plotted against the inhibitor concentration with the
576 program Origin2018 (https://www.originlab.com). All measurements were made in triplicate. The K_i values
577 were calculated using Eq. 2. For measurements using the TECAN Infinite F2000 PRO plate reader each well
578 contained 200 μL, composed of 180 μL buffer, 5 μL enzyme in buffer, 10 μL inhibitor in DMSO or ethanol,
579 and 5 μL substrate in DMSO (measuring conditions for all the proteases are summarized in Table S1,
580 supplementary information). The amount of solvent in these experiments was 7.5%. The reaction was
581 monitored for 10 min, fluorescence readout was performed in 30 s intervals. All measurements were made
582 in triplicate. IC_{50} values were calculated using GRAFIT (Version 6.0.12; Erithacus Software Limited, East
583 Grinstead, West Sussex, UK) by fitting the remaining enzymatic activity to the four-parameter IC_{50} equation
584 (Eq. 1):[47,48]

585

586 $$Y = \frac{Y_{max} - Y_{min}}{1 + \left(\frac{[I]}{IC_{50}}\right)^{s}} \hspace{4cm} \text{(Eq. 1)}$$

587

588 with Y as the substrate hydrolysis rate obtained as fluorescence increase over time (ΔF/min), Y_{max} as
589 maximum value of the dose–response curve, measured at inhibitor concentrations of [I] = 0 μM, Y_{min} as the
590 minimum value of the dose–response curve, obtained at high inhibitor concentrations, and s as the Hill
591 coefficient. The K_i value was calculated using Eq. 2:[48]

592

$$K_i = \frac{IC_{50}}{1+\frac{[S]}{K_m}} \qquad\qquad\qquad\text{(Eq. 2)}$$

594

595   with [S] being the used substrate concentration and K_m as the substrate concentration reaching half maximal
596   hydrolysis activity (determined in a separate experiment).
597   **FGA146**, the indole harboring compound, was the only one revealing a strong fluorescence in this
598   assay at higher concentrations. To rule out that bleaching of this fluorescence interferes with our readout by
599   overlaying the fluorescence increase caused by the enzymatic substrate cleavage, control measurements
600   were performed. Therefore, the assay was repeated without addition of substrate, hence, the negative slope
601   due to bleaching of **FGA146** was determined. The relative activity values were then corrected by the negative
602   slope of each inhibitor concentration (Fig. S1).

603

604   **Cell-based antiviral activity and cytotoxicity assays.** Huh-7 cells that overexpress human angiotensin-
605   converting enzyme 2 (ACE2) (Huh-7-ACE2; kindly provided by Friedemann Weber (Institute of Virology, Justus
606   Liebig University Giessen)) were grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with
607   10% fetal bovine serum (FBS) and antibiotics (100 U/mL penicillin and 100 µg/mL streptomycin) at 37 °C in
608   an atmosphere containing 5% CO_2. The SARS-CoV-2 isolate Munich 929[28] was kindly provided by Christian
609   Drosten (Institute of Virology, Charité-Universitätsmedizin, Berlin). Cytotoxic concentrations 50% (CC_{50}) of
610   the compounds used in antiviral activity assays were determined using MTT assays as described previously[49].
611   To determine effective concentrations 50% (EC_{50}) of the respective compounds, Huh-7-ACE2 cells were
612   inoculated with SARS-CoV-2 at a multiplicity of infection (MOI) of 0.1 plaque-forming units (pfu) per cell. After
613   incubation for 1 h at 33 °C, the virus inoculum was replaced with fresh cell culture medium containing the
614   test compounds at the indicated concentration. After 23 h at 33 °C, the cell culture supernatants were
615   collected and virus titers were determined by virus plaque assay as described previously[50].

616

617   **Crystallization data collection and structure determination.** Crystallization trials were performed at 295 K
618   using the sitting-drop vapor-diffusion method with commercial screening solutions including JBScreen Classic
619   and Wizard I–IV (Jena Bioscience, Jena, Germany) in 96-well sitting-drop plates (Swissci MRC; Molecular
620   Dimensions, Suffolk, England). Drops were set up by mixing equal volumes (0.2 µL) of protein-containing
621   solution (8 mg/mL) and reservoir solution using a Cartesian Honeybee System (Genomic Solutions, Irvine,
622   USA) nano-dispenser robot and equilibrated against 50 µL reservoir solution. Both crystals of the apo form
623   of the protein were obtained in 0.1 M sodium HEPES at pH 7.0 containing 22% PEG 4000 and 3% DMSO. The
624   complex with compound **FGA146** gave single well-diffracting crystals that were obtained in 0.1 M Bis-TRIS at
625   pH 6.5 containing 18% PEG 3350, and the complex with compound **FGA147** gave single well-diffracting
626   crystals that were obtained in 0.1 M TRIS-HCl at pH 8.5 containing 20% PEG 2000 MME and 10 mM $NiCl_2$.
627   For data collection, crystals were cryo-protected with a cryo-solution containing the reservoir
628   supplemented with 30 % (v/v) glycerol and flash-cooled in liquid nitrogen. X-Ray data collection experiments
629   were performed at the ALBA Synchrotron (Cerdanyola del Vallès, Spain) BL13 XALOC beamline, and at the
630   ESRF Synchrotron (Grenoble, France) ID30B beamline. Data were indexed and integrated, scaled and merged
631   using XDS[51]. The structures were solved by molecular replacement using the previously reported SARS-CoV2
632   M^{pro} structure (PDB: 7K3T) with Molrep[52]. The initial model was first refined using Refmac5[53] and alternating
633   manual building with Coot[54]. MolRep and Refmac5 are part of the CCP4 suite[55]. The final model was obtained
634   by repetitive cycles of refinement; solvent molecules were added automatically and inspected visually for
635   chemically plausible positions. The inhibitor molecule was added manually. The stereochemical quality of the
636   model was assessed with MolProbity[56]. The structural figures were generated using the Pymol program
637   (http://www.pymol.org). Data processing and refinement statistics are listed in Table S2 (Supplementary
638   information).

639

640   **Circular dichroism.** Circular dichroism measurements were carried out on a JASCO J-720 (Jasco, Tokyo, Japan)
641   spectropolarimeter equipped with a Peltier type temperature controller and a thermostatized cuvette cell
642   linked to a thermostatic bath. Spectra were recorded in 0.1 cm path length quartz cells with a response time

23

of 4 sec and a band width of 2 nm. The protein concentration used was 0.15 mg/mL in 20 mM Tris-HCl buffer at pH 7.5 and 100 mM NaCl. The observed ellipticities were converted into the molar ellipticities [θ] based on a mean molecular mass per residue of 110.45 Da. Thermal denaturation experiments were performed by increasing the temperature from 20 to 80 °C at 30 °C/hour. Tm represents the temperature at the midpoint of the unfolding transition. The CD signal was followed at 230 nm and the concentration of organic solvent was kept constant at 2.5%. Two concentrations of each compound were used (25 and 100 μM).

QM/MM simulations.

After setting up the molecular models, an additive hybrid QM/MM scheme was employed for constructing of the total Hamiltonian, where the total energy is obtained as a sum of different terms:

$$E_{QM/MM} = E_{QM} + E_{QM/MM}^{elect} + E_{QM/MM}^{vdW} + E_{MM} \qquad \text{(Eq. 3)}$$

In Equation (3), E_{QM} describes the atoms in the QM region, $E_{QM/MM}$ defines the interaction between the QM and MM region (both electrostatic and dispersion terms), and E_{MM} describes the rest of the MM region. The QM subset of atoms includes the P1' and P1 positions of the inhibitor, together with C145 and H41 residues of the protein. Four link atoms were inserted where the QM/MM boundary intersected covalent bonds in the positions indicated on Figure S9. Thus, QM part consisted of 57 atoms for both inhibitors. All the calculations were performed with the QMCube suite,[57] for which the combination of the OpenMM and Gaussian09[27] programs was used for constructing the potential energy function. The AMBER ff03[58] and the TIP3P[21] force fields were selected to describe the MM atoms, and the Minnesota functional M06-2X[59] with the split-valence 6-31+G(d,p) basis set[60] were used to treat the QM subset of atoms. This functional has been tested and shown to be suitable for modelling this type of reactivity.[24,25,30,31,61-64] The position of any atom over 25 Å from the substrate was fixed to speed up the calculations.

Reaction mechanisms for each inhibitor were initially explored using the nudged elastic band (NEB)[65] approach to set up plausible starting geometries for the transition structures. Then, they were localized and characterized by a micro-macro[66,67] Hessian-based localization scheme, and minimum energy paths (MEP) were traced towards the corresponding minima. The information obtained in this stage was used in the fine-tuning of the calculation of the free energy surface, in terms of potential of mean force (PMF). The PMF for each chemical step was obtained using the combination of the umbrella sampling (US) approach[68] with the weighted histogram analysis method (WHAM).[69] Series of MD simulations were performed adding a restraint along the collective reaction coordinate *s*, with an umbrella force constant of 3000 kJ·mol^{-1}·Å^{-2}. In every window, QM/MM MD-NVT simulations were performed with a total of 4.75 ps at 310 K with a time step of 0.5 fs (a total of 9500 steps). The definition of the *s* coordinate has been always reduced to a combination of distances. Thus, for both **FGA146** and **FGA147** inhibitors we considered the same inhibition mechanism, and the following internal coordinates were included in the s coordinate: d(Sγ, C$_{19}$), d(Sγ, Hγ), d(Hγ, Nε) and d(Hγ, C$_{20}$). All the information needed to define the equally distributed milestones from which the collective variable *s* is constructed were obtained from the analysis of the different MEPs previously traced.

Finally, the interaction energy was computed as a contribution of each residue of the protein to the interaction energy with the QM part of the substrate was computed using the following expression:

$$E_{QM/MM}^{Int} = \sum \left\langle \Psi \left| \frac{q_{MM}}{r_{e,MM}} \right| \Psi \right\rangle + \sum \sum \frac{Z_{QM} q_{MM}}{r_{QM,MM}} + E_{QM/MM}^{vdW} \qquad \text{(Eq. 4)}$$

This interaction energy can be exactly decomposed in a sum over residues provided that the polarized wave function (Ψ) is employed to evaluate this energy contribution. Because of the large number of structures that must be evaluated to obtain a representative population, the QM sub-set of atoms were described by the semiempirical Hamiltonian AM1[70] in these QM/MM MD calculations.

Data availability

693   The atomic coordinates and structure factors have been deposited into the Protein Data Bank with accession
694   codes 8BGA and 8BGD. All data are available from the corresponding author upon reasonable request. Source
695   data are provided with this paper.
696

References

(1) Organization, W. H. Weekly epidemiological update on COVID-1919 - 21 December 2022. https://www.who.int/publications/m/item/covid-19-weekly-epidemiological-update---21-december-2022.

(2) Editorial "There's no room for COVID complacency in 2023". *Nature* **2023**, *613*, 7.

(3) Cameroni, E.; Bowen, J. E.; Rosen, L. E.; Saliba, C.; Zepeda, S. K.; Culap, K.; Pinto, D.; VanBlargan, L. A.; De Marco, A.; di Iulio, J.; Zatta, F.; Kaiser, H.; Noack, J.; Farhat, N.; Czudnochowski, N.; Havenar-Daughton, C.; Sprouse, K. R.; Dillen, J. R.; Powell, A. E.; Chen, A.; Maher, C.; Yin, L.; Sun, D.; Soriaga, L.; Bassi, J.; Silacci-Fregni, C.; Gustafsson, C.; Franko, N. M.; Logue, J.; Iqbal, N. T.; Mazzitelli, I.; Geffner, J.; Grifantini, R.; Chu, H.; Gori, A.; Riva, A.; Giannini, O.; Ceschi, A.; Ferrari, P.; Cippa, P. E.; Franzetti-Pellanda, A.; Garzoni, C.; Halfmann, P. J.; Kawaoka, Y.; Hebner, C.; Purcell, L. A.; Piccoli, L.; Pizzuto, M. S.; Walls, A. C.; Diamond, M. S.; Telenti, A.; Virgin, H. W.; Lanzavecchia, A.; Snell, G.; Veesler, D.; Corti, D. Broadly neutralizing antibodies overcome SARS-CoV-2 Omicron antigenic shift. *Nature* **2022,** *602* (7898), 664-670.

(4) Cao, Y.; Wang, J.; Jian, F.; Xiao, T.; Song, W.; Yisimayi, A.; Huang, W.; Li, Q.; Wang, P.; An, R.; Wang, J.; Wang, Y.; Niu, X.; Yang, S.; Liang, H.; Sun, H.; Li, T.; Yu, Y.; Cui, Q.; Liu, S.; Yang, X.; Du, S.; Zhang, Z.; Hao, X.; Shao, F.; Jin, R.; Wang, X.; Xiao, J.; Wang, Y.; Xie, X. S. Omicron escapes the majority of existing SARS-CoV-2 neutralizing antibodies. *Nature* **2022,** *602* (7898), 657-663.

(5) Gottlieb, R. L.; Vaca, C. E.; Paredes, R.; Mera, J.; Webb, B. J.; Perez, G.; Oguchi, G.; Ryan, P.; Nielsen, B. U.; Brown, M.; Hidalgo, A.; Sachdeva, Y.; Mittal, S.; Osiyemi, O.; Skarbinski, J.; Juneja, K.; Hyland, R. H.; Osinusi, A.; Chen, S.; Camus, G.; Abdelghany, M.; Davies, S.; Behenna-Renton, N.; Duff, F.; Marty, F. M.; Katz, M. J.; Ginde, A. A.; Brown, S. M.; Schiffer, J. T.; Hill, J. A.; Investigators, G.-U.-. Early Remdesivir to Prevent Progression to Severe Covid-19 in Outpatients. *N Engl J Med* **2022**, *386* (4), 305-315.

(6) Jayk Bernal, A.; Gomes da Silva, M. M.; Musungaie, D. B.; Kovalchuk, E.; Gonzalez, A.; Delos Reyes, V.; Martin-Quiros, A.; Caraco, Y.; Williams-Diaz, A.; Brown, M. L.; Du, J.; Pedley, A.; Assaid, C.; Strizki, J.; Grobler, J. A.; Shamsuddin, H. H.; Tipping, R.; Wan, H.; Paschke, A.; Butterton, J. R.; Johnson, M. G.; De Anda, C.; Group, M. O.-O. S. Molnupiravir for Oral Treatment of Covid-19 in Nonhospitalized Patients. *N Engl J Med* **2022**, *386* (6), 509-520.

(7) Dryden-Peterson, S.; Kim, A.; Kim, A. Y.; Caniglia, E. C.; Lennes, I. T.; Patel, R.; Gainer, L.; Dutton, L.; Donahue, E.; Gandhi, R. T.; Baden, L. R.; Woolley, A. E. Nirmatrelvir Plus Ritonavir for Early COVID-19 in a Large U.S. Health System : A Population-Based Cohort Study. *Ann Intern Med* **2022,** *M22*, 2141.

(8) Wu, F.; Zhao, S.; Yu, B.; Chen, Y. M.; Wang, W.; Song, Z. G.; Hu, Y.; Tao, Z. W.; Tian, J. H.; Pei, Y. Y.; Yuan, M. L.; Zhang, Y. L.; Dai, F. H.; Liu, Y.; Wang, Q. M.; Zheng, J. J.; Xu, L.; Holmes, E. C.; Zhang, Y. Z. A new coronavirus associated with human respiratory disease in China. *Nature* **2020,** *579* (7798), 265-269.

(9) Yan, S.; Wu, G. Potential 3-chymotrypsin-like cysteine protease cleavage sites in the coronavirus polyproteins pp1a and pp1ab and their possible relevance to COVID-19 vaccine and drug development. *Faseb j* **2021,** *35* (5), e21573.

(10) Yang, H.; Yang, M.; Ding, Y.; Liu, Y.; Lou, Z.; Zhou, Z.; Sun, L.; Mo, L.; Ye, S.; Pang, H.; Gao, G. F.; Anand, K.; Bartlam, M.; Hilgenfeld, R.; Rao, Z. The crystal structures of severe acute respiratory syndrome virus main protease and its complex with an inhibitor. *Proc Natl Acad Sci U S A* **2003,** *100* (23), 13190-5.

(11) Anand, K.; Palm, G. J.; Mesters, J. R.; Siddell, S. G.; Ziebuhr, J.; Hilgenfeld, R. Structure of coronavirus main proteinase reveals combination of a chymotrypsin fold with an extra alpha-helical domain. *EMBO J* **2002,** *21* (13), 3213-24.

(12) Anand, K.; Ziebuhr, J.; Wadhwani, P.; Mesters, J. R.; Hilgenfeld, R. Coronavirus main proteinase (3CLpro) structure: basis for design of anti-SARS drugs. *Science* **2003,** *300* (5626), 1763-7.

(13) Hayden, F. G.; Turner, R. B.; Gwaltney, J. M.; Chi-Burris, K.; Gersten, M.; Hsyu, P.; Patick, A. K.; Smith, G. J., 3rd; Zalman, L. S. Phase II, randomized, double-blind, placebo-controlled studies of ruprintrivir nasal spray 2-percent suspension for prevention and treatment of experimentally induced rhinovirus colds in healthy volunteers. *Antimicrob Agents Chemother* **2003,** *47* (12), 3907-16.

(14) Kim, Y.; Liu, H.; Galasiti Kankanamalage, A. C.; Weerasekara, S.; Hua, D. H.; Groutas, W. C.; Chang, K. O.; Pedersen, N. C. Reversal of the Progression of Fatal Coronavirus Infection in Cats by a Broad-Spectrum Coronavirus Protease Inhibitor. *PLoS Pathog* **2016,** *12* (3), e1005531.

(15) Yang, H.; Xie, W.; Xue, X.; Yang, K.; Ma, J.; Liang, W.; Zhao, Q.; Zhou, Z.; Pei, D.; Ziebuhr, J.; Hilgenfeld, R.; Yuen, K. Y.; Wong, L.; Gao, G.; Chen, S.; Chen, Z.; Ma, D.; Bartlam, M.; Rao, Z. Design of wide-spectrum inhibitors targeting coronavirus main proteases. *PLoS Biol* **2005,** *3* (10), e324.

750 (16) Xiong, M.; Su, H.; Zhao, W.; Xie, H.; Shao, Q.; Xu, Y. What coronavirus 3C-like protease tells us: From
751 structure, substrate selectivity, to inhibitor design. *Med Res Rev* **2021,** *41* (4), 1965-1998.
752 (17) Agost-Beltran, L.; de la Hoz-Rodriguez, S.; Bou-Iserte, L.; Rodriguez, S.; Fernandez-de-la-Pradilla, A.;
753 Gonzalez, F. V. Advances in the Development of SARS-CoV-2 Mpro Inhibitors. *Molecules* **2022,** *27* (8).
754 (18) Dai, W.; Zhang, B.; Jiang, X. M.; Su, H.; Li, J.; Zhao, Y.; Xie, X.; Jin, Z.; Peng, J.; Liu, F.; Li, C.; Li, Y.; Bai, F.;
755 Wang, H.; Cheng, X.; Cen, X.; Hu, S.; Yang, X.; Wang, J.; Liu, X.; Xiao, G.; Jiang, H.; Rao, Z.; Zhang, L. K.; Xu, Y.;
756 Yang, H.; Liu, H.
757 Structure-based design of antiviral drug candidates targeting the SARS-CoV-2 main protease. *Science* **2020,**
758 *368* (6497), 1331-1335.
759 (19) Zhang, L.; Lin, D.; Sun, X.; Curth, U.; Drosten, C.; Sauerhering, L.; Becker, S.; Rox, K.; Hilgenfeld, R. Crystal
760 structure of SARS-CoV-2 main protease provides a basis for design of improved alpha-ketoamide inhibitors.
761 *Science* **2020,** *368* (6489), 409-412.
762 (20) Hoffman, R. L.; Kania, R. S.; Brothers, M. A.; Davies, J. F.; Ferre, R. A.; Gajiwala, K. S.; He, M.; Hogan, R.
763 J.; Kozminski, K.; Li, L. Y.; Lockner, J. W.; Lou, J.; Marra, M. T.; Mitchell, L. J., Jr.; Murray, B. W.; Nieman, J. A.;
764 Noell, S.; Planken, S. P.; Rowe, T.; Ryan, K.; Smith, G. J., 3rd; Solowiej, J. E.; Steppan, C. M.; Taggart, B.
765 Discovery of Ketone-Based Covalent Inhibitors of Coronavirus 3CL Proteases for the Potential Therapeutic
766 Treatment of COVID-19. *J Med Chem* **2020,** *63* (21), 12725-12747.
767 (21) Owen, D. R.; Allerton, C. M. N.; Anderson, A. S.; Aschenbrenner, L.; Avery, M.; Berritt, S.; Boras, B.; Cardin,
768 R. D.; Carlo, A.; Coffman, K. J.; Dantonio, A.; Di, L.; Eng, H.; Ferre, R.; Gajiwala, K. S.; Gibson, S. A.; Greasley, S.
769 E.; Hurst, B. L.; Kadar, E. P.; Kalgutkar, A. S.; Lee, J. C.; Lee, J.; Liu, W.; Mason, S. W.; Noell, S.; Novak, J. J.;
770 Obach, R. S.; Ogilvie, K.; Patel, N. C.; Pettersson, M.; Rai, D. K.; Reese, M. R.; Sammons, M. F.; Sathish, J. G.;
771 Singh, R. S. P.; Steppan, C. M.; Stewart, A. E.; Tuttle, J. B.; Updyke, L.; Verhoest, P. R.; Wei, L.; Yang, Q.; Zhu,
772 Y. An oral SARS-CoV-2 M(pro) inhibitor clinical candidate for the treatment of COVID-19. *Science* **2021,** *374*
773 (6575), 1586-1593.
774 (22) Jin, Z.; Du, X.; Xu, Y.; Deng, Y.; Liu, M.; Zhao, Y.; Zhang, B.; Li, X.; Zhang, L.; Peng, C.; Duan, Y.; Yu, J.;
775 Wang, L.; Yang, K.; Liu, F.; Jiang, R.; Yang, X.; You, T.; Liu, X.; Yang, X.; Bai, F.; Liu, H.; Liu, X.; Guddat, L. W.;
776 Xu, W.; Xiao, G.; Qin, C.; Shi, Z.; Jiang, H.; Rao, Z.; Yang, H. Structure of M(pro) from SARS-CoV-2 and
777 discovery of its inhibitors. *Nature* **2020,** *582* (7811), 289-293.
778 (23) Latorre, A.; Schirmeister, T.; Kesselring, J.; Jung, S.; Johé, P.; Hellmich, U. A.; Heilos, A.; Engels, B.;
779 Krauth-Siegel, R. L.; Dirdjaja, N.; Bou-Iserte, L.; Rodríguez, S.; González, F. V. Dipeptidyl Nitroalkenes as
780 Potent Reversible Inhibitors of Cysteine Proteases Rhodesain and Cruzain. *ACS Med Chem Lett* **2016,** *7* (12),
781 1073-1076.
782 (24) Arafet, K.; González, F. V.; Moliner, V. Quantum Mechanics/Molecular Mechanics Studies of the
783 Mechanism of Cysteine Proteases Inhibition by Dipeptidyl Nitroalkenes. *Chemistry* **2020,** *26* (9), 2002-2012.
784 (25) Arafet, K.; Serrano-Aparicio, N.; Lodola, A.; Mulholland, A. J.; González, F. V.; Świderek, K.; Moliner, V.
785 Mechanism of inhibition of SARS-CoV-2 M(pro) by N3 peptidyl Michael acceptor explained by QM/MM
786 simulations and design of new derivatives with tunable chemical reactivity. *Chem Sci* **2020,** *12* (4), 1433-1444.
787 (26) Mellott, D. M.; Tseng, C. T.; Drelich, A.; Fajtová, P.; Chenna, B. C.; Kostomiris, D. H.; Hsu, J.; Zhu, J.;
788 Taylor, Z. W.; Kocurek, K. I.; Tat, V.; Katzfuss, A.; Li, L.; Giardini, M. A.; Skinner, D.; Hirata, K.; Yoon, M. C.;
789 Beck, S.; Carlin, A. F.; Clark, A. E.; Beretta, L.; Maneval, D.; Hook, V.; Frueh, F.; Hurst, B. L.; Wang, H.;
790 Raushel, F. M.; O'Donoghue, A. J.; de Siqueira-Neto, J. L.; Meek, T. D.; McKerrow, J. H. A Clinical-Stage
791 Cysteine Protease Inhibitor blocks SARS-CoV-2 Infection of Human and Monkey Cells. *ACS Chem Biol* **2021,**
792 *16* (4), 642-650.
793 (27) Shang, J.; Wan, Y.; Luo, C.; Ye, G.; Geng, Q.; Auerbach, A.; Li, F. Cell entry mechanisms of SARS-CoV-2.
794 *Proc Natl Acad Sci U S A* **2020,** *117* (21), 11727-11734.
795 (28) Hoffmann, M.; Kleine-Weber, H.; Schroeder, S.; Krüger, N.; Herrler, T.; Erichsen, S.; Schiergens, T. S.;
796 Herrler, G.; Wu, N. H.; Nitsche, A.; Müller, M. A.; Drosten, C.; Pöhlmann, S. SARS-CoV-2 Cell Entry Depends
797 on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* **2020,** *181* (2), 271-280.e8.
798 (29) Liu, T.; Luo, S.; Libby, P.; Shi, G. P. Cathepsin L-selective inhibitors: A potentially promising treatment for
799 COVID-19 patients. *Pharmacol Ther* **2020,** *213*, 107587.
800 (30) Świderek, K.; Moliner, V. Revealing the molecular mechanisms of proteolysis of SARS-CoV-2 M(pro) by
801 QM/MM computational methods. *Chem Sci* **2020,** *11* (39), 10626-10630.

(31) Marti, S.; Arafet, K.; Lodola, A.; Mulholland, A. J.; Swiderek, K.; Moliner, V. Impact of Warhead Modulations on the Covalent Inhibition of SARS-CoV-2 M(pro) Explored by QM/MM Simulations. *ACS Catal* **2022,** *12* (1), 698-708.

(32) Zhang, L.; Lin, D.; Kusov, Y.; Nian, Y.; Ma, Q.; Wang, J.; von Brunn, A.; Leyssen, P.; Lanko, K.; Neyts, J.; de Wilde, A.; Snijder, E. J.; Liu, H.; Hilgenfeld, R. α-Ketoamides as Broad-Spectrum Inhibitors of Coronavirus and Enterovirus Replication: Structure-Based Design, Synthesis, and Activity Assessment. *J Med Chem* **2020,** *63* (9), 4562-4578.

(33) Jaskolski, M.; Dauter, Z.; Shabalin, I. G.; Gilski, M.; Brzezinski, D.; Kowiel, M.; Rupp, B.; Wlodawer, A. Crystallographic models of SARS-CoV-2 3CL(pro): in-depth assessment of structure quality and validation. *IUCrJ* **2021,** *8* (Pt 2), 238-256.

(34) Günther, S.; Reinke, P. Y. A.; Fernández-García, Y.; Lieske, J.; Lane, T. J.; Ginn, H. M.; Koua, F. H. M.; Ehrt, C.; Ewert, W.; Oberthuer, D.; Yefanov, O.; Meier, S.; Lorenzen, K.; Krichel, B.; Kopicki, J.-D.; Gelisio, L.; Brehm, W.; Dunkel, I.; Seychell, B.; Gieseler, H.; Norton-Baker, B.; Escudero-Pérez, B.; Domaracky, M.; Saouane, S.; Tolstikova, A.; White, T. A.; Hänle, A.; Groessler, M.; Fleckenstein, H.; Trost, F.; Galchenkova, M.; Gevorkov, Y.; Li, C.; Awel, S.; Peck, A.; Barthelmess, M.; Schlünzen, F.; Lourdu Xavier, P.; Werner, N.; Andaleeb, H.; Ullah, N.; Falke, S.; Srinivasan, V.; França, B. A.; Schwinzer, M.; Brognaro, H.; Rogers, C.; Melo, D.; Zaitseva-Doyle, J. J.; Knoska, J.; Peña-Murillo, G. E.; Mashhour, A. R.; Hennicke, V.; Fischer, P.; Hakanpää, J.; Meyer, J.; Gribbon, P.; Ellinger, B.; Kuzikov, M.; Wolf, M.; Beccari, A. R.; Bourenkov, G.; von Stetten, D.; Pompidor, G.; Bento, I.; Panneerselvam, S.; Karpics, I.; Schneider, T. R.; Garcia-Alai, M. M.; Niebling, S.; Günther, C.; Schmidt, C.; Schubert, R.; Han, H.; Boger, J.; Monteiro, D. C. F.; Zhang, L.; Sun, X.; Pletzer-Zelgert, J.; Wollenhaupt, J.; Feiler, C. G.; Weiss, M. S.; Schulz, E.-C.; Mehrabi, P.; Karničar, K.; Usenik, A.; Loboda, J.; Tidow, H.; Chari, A.; Hilgenfeld, R.; Uetrecht, C.; Cox, R.; Zaliani, A.; Beck, T.; Rarey, M.; Günther, S.; Turk, D.; Hinrichs, W.; Chapman, H. N.; Pearson, A. R.; Betzel, C.; Meents, A. X-ray screening identifies active site and allosteric inhibitors of SARS-CoV-2 main protease. *Science* **2021**, *372*, 642-646.

(35) Kovalevsky, A.; Coates, L.; Kneller, D. W.; Ghirlando, R.; Aniana, A.; Nashed, N. T.; Louis, J. M. Unmasking the Conformational Stability and Inhibitor Binding to SARS-CoV-2 Main Protease Active Site Mutants and Miniprecursor. *J. Mol. Biol.* **2022**, *434*, 167876-167891.

(36) Lu, J.; Chen, S. A.; Khan, M. B.; Brassard, R.; Arutyunova, E.; Lamer, T.; Vuong, W.; Fischer, C.; Young, H. S.; Vederas, J. C.; Lemieux, M. J. Crystallization of Feline Coronavirus M^pro With GC376 Reveals Mechanism of Inhibition. *Frontiers in Chemistry* **2022**, *1*0, A852210.

(37) Nashed, N. T.; Kneller, D. W.; Coates, L.; Ghirlando, R.; Aniana, A.; Kovalevsky, A.; Louis, J. M. Autoprocessing and oxyanion loop reorganization upon GC373 and nirmatrelvir binding of monomeric SARS-CoV-2 main protease catalytic domain. *Communications Biology* **2022**, *5*, 976-990.

(38) Sacco, M. D.; Hu, Y.; Gongora, M. V.; Meilleur, F.; Kemp, M. T.; Zhang, X.; Wang, J.; Chen, Y. The P132H mutation in the main protease of Omicron SARS-CoV-2 decreases thermal stability without compromising catalysis or small-molecule drug inhibition. *Cell Research* **2022**, *32*, 498-500.

(39) Wossner, N.; Alhalabi, Z.; Gonzalez, J.; Swyter, S.; Gan, J.; Schmidtkunz, K.; Zhang, L.; Vaquero, A.; Ovaa, H.; Einsle, O.; Sippl, W.; Jung, M. Sirtuin 1 Inhibiting Thiocyanates (S1th)-A New Class of Isotype Selective Inhibitors of NAD(+) Dependent Lysine Deacetylases. *Front Oncol* **2020,** *10*, 657.

(40) Simeonov, A. Recent developments in the use of differential scanning fluorometry in protein and small molecule discovery and characterization. *Expert Opin Drug Discov* **2013,** *8* (9), 1071-82.

(41) Weber, P. C.; Salemme, F. R. Applications of calorimetric methods to drug discovery and the study of protein interactions. *Curr Opin Struct Biol* **2003,** *13* (1), 115-21.

(42) Verschueren, K. H.; Pumpor, K.; Anemüller, S.; Chen, S.; Mesters, J. R.; Hilgenfeld, R. A structural view of the inactivation of the SARS coronavirus main proteinase by benzotriazole esters. *Chem Biol* **2008,** *15* (6), 597-606.

(43) Amendola, G. et al. Lead Discovery of SARS-CoV-2 Main Protease Inhibitors through Covalent Docking Based Virtual Screening. *J Chem Inf Model* **61**, 2062-2073 (2021).

(44) Desilets, A., Longpre, J.M., Beaulieu, M.E. & Leduc, R. Inhibition of human matriptase by eglin c variants. *FEBS Lett* **580**, 2227-32 (2006).

(45) Phoo, W.W. et al. Structure of the NS2B-NS3 protease from Zika virus after self-cleavage. *Nat Commun* **7**, 13410 (2016).

854 (46) Klein, P. et al. Naphthoquinones as Covalent Reversible Inhibitors of Cysteine Proteases-Studies on
855 Inhibition Mechanism and Kinetics. *Molecules* **25**(2020).
856 (47) Leatherbarrow, R.J. GraFit 6. Erithacus Software Limited: East Gristead, West Sussex. (2007).
857 (48) Ludewig, S., Kossner, M., Schiller, M., Baumann, K. & Schirmeister, T. Enzyme kinetics and hit validation
858 in fluorimetric protease assays. *Curr Top Med Chem* **10**, 368-82 (2010).
859 (49) Müller, C. et al. Broad-spectrum antiviral activity of the eIF4A inhibitor silvestrol against corona- and
860 picornaviruses. *Antiviral Res* **150**, 123-129 (2018).
861 (50) Kabsch, W. XDS. *Acta Crystallogr D Biol Crystallogr* **66**, 125-32 (2010).
862 (51) Vagin, A. & Teplyakov, A. Molecular replacement with MOLREP. *Acta Crystallogr D Biol Crystallogr* **66**,
863 22-5 (2010).
864 (53) Murshudov, G.N., Vagin, A.A. & Dodson, E.J. Refinement of macromolecular structures by the
865 maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* **53**, 240-55 (1997).
866 (54) Emsley, P., Lohkamp, B., Scott, W.G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr*
867 *D Biol Crystallogr* **66**, 486-501 (2010).
868 (55) Winn, M.D. et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol*
869 *Crystallogr* **67**, 235-42 (2011).
870 (56) Chen, V.B. et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta*
871 *Crystallogr D Biol Crystallogr* **66**, 12-21 (2010).
872 (57) Martí, S. QMCube (QM3): An All-purpose Suite for Multiscale QM/MM Calculations. *Journal of*
873 *Computational Chemistry* **42**, 447-457 (2021).
874 (58) Duan, Y. et al. A Point-charge Force Field for Molecular Mechanics Simulations of Proteins Based on
875 Condensed-phase Quantum Mechanical Calculations. *Journal of Computational Chemistry* **24**, 1999-2012
876 (2003).
877 (59) Zhao, Y. & Truhlar, D.G. The M06 Suite of Density Functionals for Main Group Thermochemistry,
878 Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: Two New
879 Functionals and Systematic Testing of Four M06-class Functionals and 12 Other Functionals. *Theoretical*
880 *Chemistry Accounts* **120**, 215-241 (2008).
881 (60) Hehre, W.J., Radom, L., Schleyer, P.V.R. & Pople, J.A. *Ab Initio Molecular Orbital Theory*, (John Wiley,
882 New York, 1986).
883 (61) Arafet, K., Ferrer, S. & Moliner, V. Computational Study of the Catalytic Mechanism of the Cruzain
884 Cysteine Protease. *ACS Catalysis* **7**, 1207-1215 (2017).
885 (62) Arafet, K., Świderek, K. & Moliner, V. Computational Study of the Michaelis Complex Formation and
886 the Effect on the Reaction Mechanism of Cruzain Cysteine Protease. *ACS Omega* **3**, 18613-18622 (2018).
887 (63) Arafet, K., González, F.V. & Moliner, V. Elucidating the Dual Mode of Action of Dipeptidyl Enoates in the
888 Inhibition of Rhodesain Cysteine Proteases. *Chemistry – A European Journal* **27**, 10142-10150
889 (2021).
890 (64) Chan, H.T.H. et al. Discovery of SARS-CoV-2 Mpro Peptide Inhibitors from Modelling Substrate and
891 Ligand Binding. *Chemical Science* **12**, 13686-13703 (2021).
892 (65) Henkelman, G. & Jónsson, H. Improved tangent estimate in the nudged elastic band method for finding
893 minimum energy paths and saddle points. *The Journal of Chemical Physics* **113**, 9978-9985 (2000).
894 (66) Turner, A.J., Moliner, V. & Williams, I.H. Transition-state Structural Refinement with GRACE and
895 CHARMM: Flexible QM/MM Modelling for Lactate Dehydrogenase. *Physical Chemistry Chemical*
896 *Physics* **1**, 1323-1331 (1999).
897 (67) Martí, S., Moliner, V. & Tuñón, I. Improving the QM/MM Description of Chemical Processes: A Dual
898 Level Strategy to Explore the Potential Energy Surface in Very Large Systems. *Journal of Chemical*
899 *Theory and Computation* **1**, 1008-1016 (2005).
900 (68) Torrie, G.M. & Valleau, J.P. Non-physical Sampling Distributions in Monte_Carlo Free-Energy
901 Estimation-Umbrella Sampling. *Journal of Computational Physics* **23**, 187-199 (1977).
902 (69) Kumar, S., Bouzida, D., Swendsen, R.H., Kollman, P.A. & Rosenberg, J.M. The Weighted Histogram
903 Analysis Method for Free-Energy Calculations on Biomolecules.1. The Method. *Journal of*
904 *Computational Chemistry* **13**, 1011-1021 (1992).

905   40.   Dewar, M.J.S., Zoebisch, E.G., Healy, E.F. & Stewart, J.J.P. The Development and Use of Quantum
906   Mechanical Molecular-models. 76. AM1: A New General Purpose Quantum Mechanical Molecular Model. *J.*
907   *Am. Chem. Soc.* **107**, 3902-3909 (1985)
908
909
910

Author contributions

All authors contributed to this work. A.R., V.M. and F.V.G. conceived and designed the study; F.J.M. obtained the crystal structures and performed the CD experiments; A.G.-M. performed the kinetic experiments; E.S. performed the cloning; S.H.-R. performed the synthesis; S.J.H. performed protein expression, purification, and enzymatic assays of M^{pro} and off-targets; C.M. and J.Z. performed the virus and cell assays; S.M., K.A. and A.L. carried out the computer simulations; All authors participated in the discussion of the results; F.J.M., F.V.G., V.M., A.L., K.Ś., A.R. and T.S. participated in writing the original draft. All authors have read and approved the published version of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SuppMat.pdf