Catalytic cleavage of HEAT and subsequent covalent binding of the tetralone moiety by the SARS-CoV-2 main protease.

Abstract: Here we present the crystal structure of SARS-CoV-2 main protease (M^{pro}) covalently bound to 2-methyl-1-tetralone. This complex was obtained by co-crystallization of M^{pro} with HEAT (2-(((4-hydroxyphenethyl)amino)methyl)-3,4-dihydronaphthalen-1(2H)-one) in the framework of a large X-ray crystallographic screening project of M^{pro} against a drug repurposing library, consisting of 5600 approved drugs or compounds in clinical phase trials. Further investigations showed that HEAT is cleaved by M^{pro} in an E1cB-like reaction mechanism into 2-methylene-1-tetralone and tyramine. The catalytic Cys145 subsequently binds covalently in a Michael addition to the methylene carbon atom of 2-methylene-1-tetralone. According to this postulated model HEAT is acting in a pro-drug-like fashion. It is metabolized by M^{pro}, followed by covalent binding of one metabolite to the active site. Protease activity assays show that this inactivates M^{pro}. The structure of the covalent adduct elucidated in this study opens up a new path for developing non-peptidic inhibitors

Introduction

The pandemic disease caused by the coronavirus SARS-CoV-2 is an international health challenge and there is an immediate need for the identification of drugs for treatment of SARS-CoV-2 infections. Infection of host cells critically depends on several molecular factors from both the host and virus (*1*, *2*). Coronaviruses are RNA-viruses with a genome of around 30 kbp. The viral gene essential for the replication of the virus is expressed as polyprotein that must be broken into functional subunits for replication and transcription activity (*1*). This proteolytic cleavage is primarily accomplished by the main protease (M^{pro}), also known as the 3C-like protease $3CL^{pro}$. M^{pro} cleaves the viral polyprotein at eleven sites. The core cleavage motif is Leu-Gln⇓(Ser/Ala/Gly) (*1*). M^{pro} possesses a chymotrypsin-like fold amended by a C-terminal helical domain and harbors a catalytic dyad comprising Cys145 and His41 (*1*). The active site is located in a cleft between the two N-terminal domains of the three domain structure of the monomer, while the C-terminal helical domain is involved in regulation of dimerization of the enzyme with a dimerization constant of 2.5 μM (*1*). Due to its central involvement in virus replication M^{pro} is well established as a prime target for antiviral compounds that can tackle coronavirus infections (*3*). Indeed first studies confirm the potential of targeting M^{pro} for inhibition of virus replication (*1*, *2*).

In order to identify further potential drug candidates with an inhibitory effect on M^{pro} we screened the "Fraunhofer IME Repurposing Collection" (*4*), consisting of 5632 individual compounds against M^{pro} by means of high-throughput X-ray crystallography. This compound collection is based on the BROAD institute repurposing library (*5*). In contrast to crystallographic screening campaigns with small molecules ("fragment screening") (*6*), our library contains compounds that are chemically more complex and of larger molecular mass. As M^{pro} crystals possess a low solvent content of only 38% (*1*) with narrow

solvent channels that could impede larger molecules from diffusing into the active site in soaking experiments, we co-crystallized M^{pro} with the individual compounds instead of soaking pre-grown M^{pro} crystals in compound-containing solutions. Crystals typically appeared within 2 days and were then submitted to single crystal X-ray structure determination (see Methods). From the resulting X-ray structures, several compounds were identified in complex with M^{pro}. Among these was BE-2254 (HEAT, 2-(((4-hydroxy-phenethyl)amino)methyl)-3,4-dihydronaphthalen-1(2H)-one), where a well resolved electron-density was observed in the active site of the M^{pro} crystal structure. This electron-density can be well modeled as 2-methyl-1-tetralone (METT) covalently bound to the sulfur atom of Cys145, thereby blocking access to the active site of M^{pro} and occupying the P1 site of the protease. 2-Methyl-1-tetralone is potentially a breakdown product of the parent compound HEAT included in the co-crystallization experiment.

HEAT was first synthesized by Hansen at Beiersdorf AG (7). It exists in a three-way equilibrium between two enantiomers between which it transitions via keto-enol tautomerism (8). Investigations by several groups found it to be a competitive α-adrenoreceptor antagonist *in vitro* and *in vivo* with lower affinity to the α-subtype and residual affinity to dopaminergic and serotonergic receptors (9).

Here we present the crystal structure of M^{pro} with 2-methyl-1-tetralone (METT) covalently bound to the active site to 1.8 Å resolution. To further support this observation and to understand the underlying mechanism of HEAT fragmentation by M^{pro} and its subsequent covalent binding to the active site, we conducted complementary native and small molecule mass spectrometry and computational docking studies. In addition we examined the inhibitory effect of HEAT on M^{pro} in a protease activity assay. On the basis of these results we propose a reaction mechanism and provide an outlook for the application of HEAT as a starting point for potential drug development to treat SARS-CoV-2 infections.

Results

The overall structure of the SARS-CoV-2 main protease (M^{pro}) with bound METT as obtained in our X-ray diffraction experiment is shown in figure 1 (table 1). It possesses the same overall structural features as previously described (1, 2). A closer view of the active site reveals clear difference electron-density for a small molecule covalently bound to the catalytic Cys145 which could be well modeled as METT (inset figure 1). The chemical structures of the parent compound HEAT and the observed cleavage product METT are illustrated in figure 2, clearly showing the relationship between the two molecules. This observation suggests METT to be a breakdown product of HEAT, with tyramine (TA) as the complementary product.
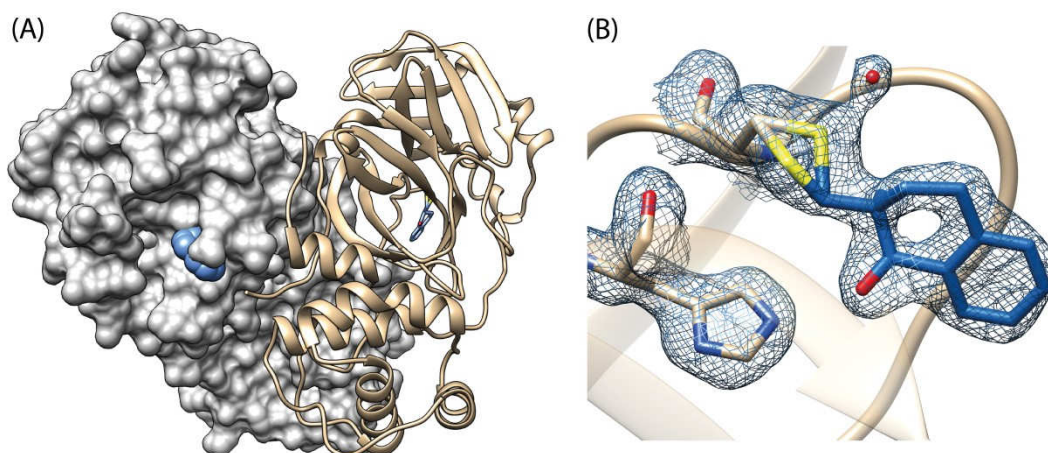
Figure 1: *X-ray structure of M^pro with covalently bound 2-methyl-1-tetralone.* (A) The biologically active form of M^pro is a homodimer found in the crystal structure. Here one monomer is depicted as surface representation, while the second monomer is shown as cartoon model. The ligand is highlighted as sphere and stick model (blue). (B) A simulated annealing omit 2Fo-Fc-map of the active site of M^pro reveals clear electron-density for a ligand covalently bound to Cys145 that can be modeled as METT. Coordinates of the ligand were removed before calculating the map (1 rmsd and carved at 1.4 Å around the atoms). Diffraction data together with coordinates were deposited in the Protein Data Bank with PDB accession code 6YNQ.
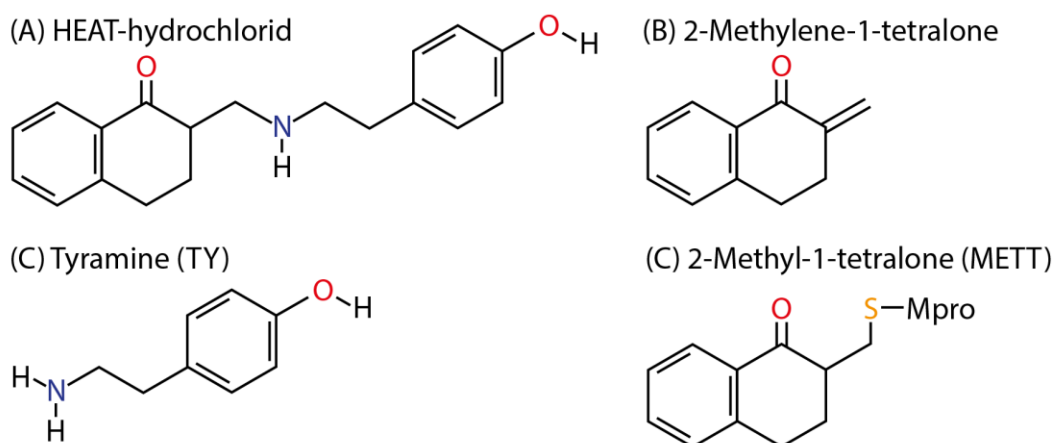

Figure 2: Chemical structures of the parent compound HEAT (A), the potential reaction intermediate 2-methylene-1-tetralone (B) and the proposed degradation products tyramine (TY, C) and 2-methyl-1-tetralone (METT) bound to M^pro (D).

To confirm the cleavage of HEAT by M^pro and subsequent covalent binding of METT to M^pro native mass spectrometry (MS) experiments were performed (*10*). In a control experiment without the addition of HEAT we observe a high intensity peak envelope corresponding to the expected mass of the M^pro dimer (67.599 +/- 5 Da, charge states +14 to +16). The monomer is also visible with

lower intensity (33.795 +/- 2 Da, charge states +9 to +11) (figure 3a). Next M^{pro} was incubated with HEAT at a molar ratio of 1:10 (M^{pro}:HEAT) overnight. After incubation of M^{pro} with HEAT two additional peak envelopes can be observed in the high mass range (figure 3b). Their assigned masses correspond to the M^{pro} dimer alone or plus one and two METT molecules (67,593/67,774/67,955 Da). The corresponding stoichiometry of M^{pro} dimer without or either one or two bound METT was approximately 4:5:1. Additionally peaks from monomeric M^{pro} bound to one METT molecule and unmodified M^{pro} are also observed. The expected ion of intact HEAT with 296 Da (charge 1+) is observed in the low mass range (suppl. fig.XX/to be added). The reaction product METT is not observed as it is too small to be detected with the instrument settings as used for these measurements.

To probe the chemical interaction between M^{pro} and HEAT, CID-MS/MS was performed by separating a molecular ion precursor (+15, dimer) in the quadrupole of the mass spectrometer and triggering its dissociation in a collision cell (see Methods). The dimer dissociated into low- and high-charge monomeric molecular ions. METT was observed to still be bound to the high-charge ions. The ratio of the dissociated monomer in the unbound to bound state was roughly 5:4. The binding of the reactive group to the dissociated molecule indicates a strong interaction between M^{pro} and METT, as would be expected for a covalent bond. Furthermore, only one METT per monomer is detected which suggests that the interaction is site-specific.
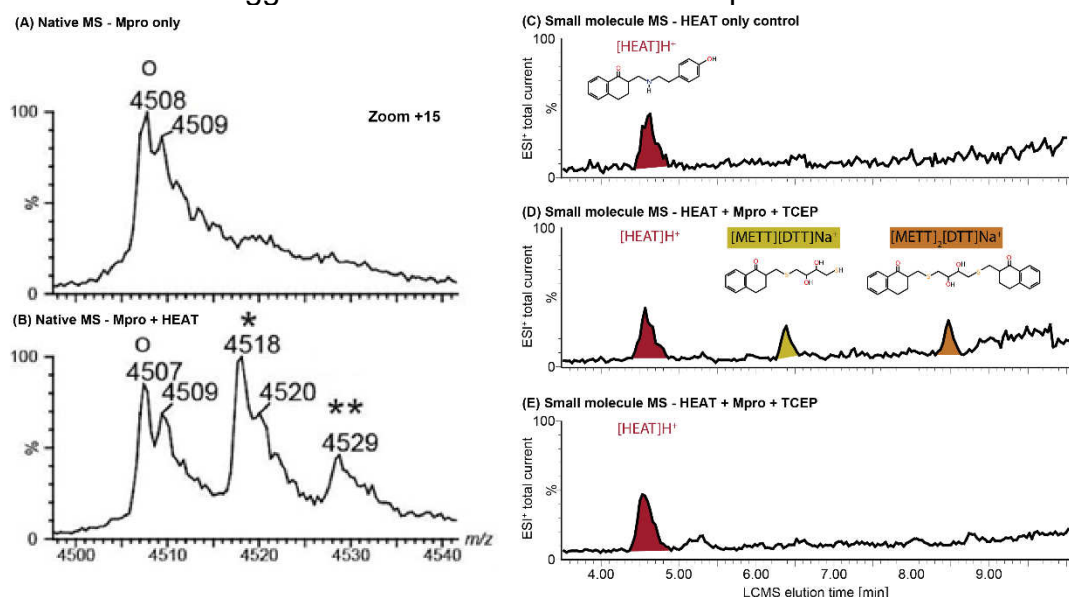

Figure 3(to be updated): *MS analysis of the educts and products of the reaction of HEAT with M^{pro}.* (A) native mass spectrum of native M^{pro} indicating the homodimeric state. (B) mass spectrum of M^{pro} after overnight incubation with HEAT leading to additional peaks with higher masses. HPLC elution profile (ESI$^+$, Total Ion Current) of the reaction products of HEAT alone (C) and with M^{pro} in the presence of DTT (D) or TCEP (E). The mass spectra corresponding to the elution peaks are shown in F-G.

In addition, the small-molecule products from the reaction of HEAT with M^{pro} were analyzed by liquid chromatography mass spectrometry (LC-MS) in the range of 100 to 1000 Da (figure 3, see Methods). Pure HEAT elutes at 4.5 min and displays major ions at 296 Da (MH$^+$) and 318 Da (MNa$^+$). After overnight incubation (12 hours) at room temperature of an excess of HEAT with M^{pro}, two new LC-MS peaks were observed at 6.4 and 8.5 min. Masses of 335.3 and 493.3 Da appear in the MS-ES+ spectrum, which could not be attributed to either of the expected reaction products METT (183 Da) or TY (137 Da). These observed masses correspond to the two possible adducts of METT and DTT (154 Da), which was used as reducing agent in the protein buffer. The observed signals correspond to MNa$^+$ ions of METT bound to one or both thiol groups of DTT ([METT][DTT]Na$^+$ or [METT]$_2$[DTT]Na$^+$, respectively). The reaction product TY was observed in the MS-ES- spectrum (137 Da, supp. figure 3) in the presence of M^{pro}. To confirm the assignment of the METT and DTT adducts we substituted DTT with tris(2-carboxyethyl)phosphine (TCEP) as reducing agent. In the absence of DTT no additional LC-MS signals besides HEAT were observed, and no masses corresponding to the METT-DTT adducts could be detected. This supports the presence of the proposed METT intermediate in the HEAT cleavage mechanism, which can be intercepted by the strong nucleophilic thiol group of DTT. In contrast, TCEP is much less nucleophilic and hence does not react with the intermediate of the HEAT cleavage.

Understanding of the underlying reaction mechanism requires also knowledge about the initial binding of HEAT to M^{pro} which was not observed in the X-ray diffraction experiment. To obtain information on the initial binding state of HEAT to M^{pro}, computational docking studies were employed using the docking tools JAMDA and HYDE (*11*) (Figure 4). We visually inspected multiple proposed non-covalent binding modes which might help to reveal the possible reaction mechanism. The keto form of HEAT (figure 4a) is found to occupy pockets P1 and P2 with a carbon-sulfur distance of 3.4 Å. HEAT forms a hydrogen bond to His163 in the active site via the carbonyl oxygen atom, a weak hydrogen bond to Cys145 via its secondary amine, and a T-stacking interaction to His41. The enol form (figure 4b) seems to be more stably bound to the active site, forming not only the two aforementioned hydrogen bonds, but also a hydrogen bond to Cys44 via its phenol moiety. Additionally, a pi-pi stacking interaction between His41 and the phenol moiety can be observed in its predicted binding mode. Moreover, the carbon-sulfur distance of 3.5 Å for the enol form of HEAT to Cys145 is well within the acceptable range for covalent bond formation. This non-covalent binding mode corresponds well to the experimentally determined binding mode of METT which occupies the S1 pocket (figure 4c). Additionally, the leaving group of the molecule is situated in the S2 pocket and is engaged in a weak pi-pi stacking interaction with His41 and a weak hydrogen bond to Cys44. We hypothesize that HEAT initially binds non-covalently to the protease and is converted to the corresponding enol form prior to the formation of the α,β-unsaturated ketone. This process might be catalyzed by residues His163 and Cys145, which are engaged in hydrogen bonds to the product METT as well as the proposed non-covalently bound keto- and enol-form of HEAT.
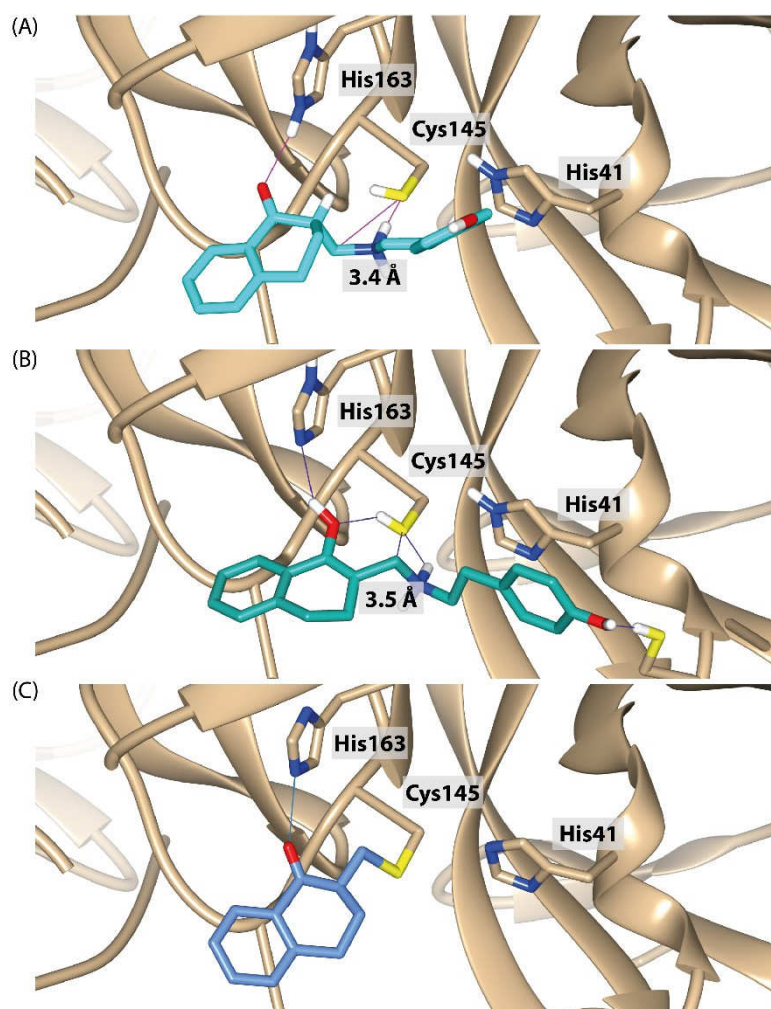
Figure 4: *Proposed non-covalent binding modes of HEAT in comparison to the experimentally determined covalent binding mode.* Hydrogen bonds are depicted as purple lines and the carbon-sulfur distance is highlighted in pink. Top: Proposed binding mode of the keto form of HEAT. Middle: Proposed binding mode of the enol form of HEAT. Bottom: Experimental binding mode of 2-methyl-1-tetralone.

Due to the covalent binding to the active-site Cys145, we expected an inhibitory effect on the enzyme function and thus virus replication. Therefore in a next step we conducted a complementary protease activity assay (*1*). At a HEAT concentration of 100 µM and 0.1 µM M^{pro}, and in the presence of 1 mM DTT no inhibitory effect of HEAT on M^{pro} was detected after 10 min incubation. This is in line with our observations from small-molecule LC-MS where DTT was found to bind to the reaction intermediate of HEAT in competition with M^{pro}. Therefore we replaced DTT by TCEP (1 mM) whilst keeping the other parameters identical. This again revealed no significant inhibition of M^{pro}.

Discussion
Combining all experimental and computational results allows us to propose a plausible, albeit highly unusual, reaction mechanism. M^{pro} is a cysteine protease possessing the conserved reactive center dyad Cys145 and His41, which together catalyze the sequence-specific hydrolysis of peptide bonds, vital for SARS-CoV-2 replication. We propose a similar mechanism for the cleavage of the drug HEAT, leading to a suicide reaction of the protease, which terminates in a covalent thioether bond between Cys145 and METT

(Figure 5). As initiation the tetralone carbonyl is activated by His163, forming an enol. The thiolate of Cys145 could serve as base, removing the hydrogen of the β-carbon and forming the enol. After enol-oxygen deprotonation by His163, an unimolecular conjugate base elimination (E1cB)-like mechanism expels the protonated tyramine amine, forming 2-methylen-1-tetralone in the active site. In a subsequent step this highly reactive intermediate reacts in a Michael addition, where the thiolate of Cys145 attacks the α,β-unsaturated 2-methylen-1-tetralone, forming a thioether bond resulting in covalently bound METT. This final reaction state was confirmed by X-ray crystallography to a resolution of 1.8 Å. As evident in our LC-MS analysis some METT can react with DTT and it is not yet clear whether this occurs within the active site or spontaneously in solution.
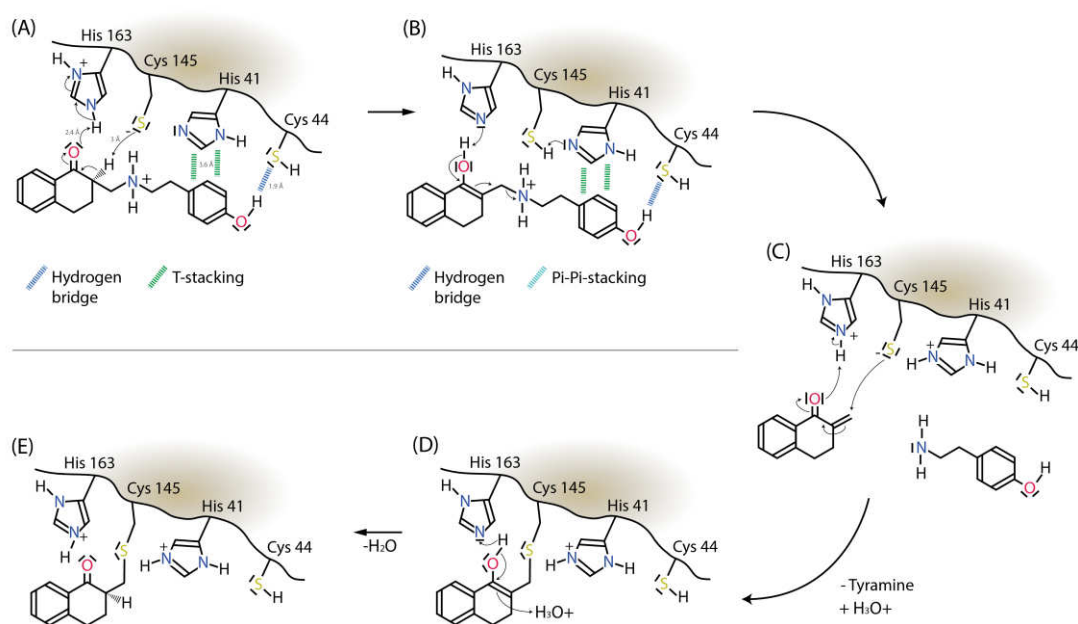

Figure 5: *Proposed reaction mechanism of HEAT with M^pro*. HEAT exists in equilibrium between a keto- (A) and enol- (B) form via tautomerism. The catalytic center bound enol form of HEAT eliminates the tyramine group, leaving 2-methylene-1-tetralone in the active site (C). This is attacked by the thiol of Cys145, forming the observed covalent thioether bond (D). This enol is again in equilibrium with the keto-form, observed in the X-ray structure (E).

The current pandemic of SARS-CoV-2 and also the earlier epidemics caused by SARS-CoV and MERS-CoV revealed the risk posed by coronaviruses and stressed the need for pan-coronavirus antiviral drugs to be better prepared for future outbreaks. M^pro is highly conserved across coronaviruses (*12*) (suppl. figure1). The active site residues His41 and Cys145 are strictly conserved. Furthermore, in our structure METT binding is stabilized by a hydrogen bond to His163 as well as by van der Waals interactions with the backbone of the oxyanion hole formed by residues Leu141 and Asn142 and neighboring Phe140. These residues are all essential for the protease activity of M^pro and

thus are conserved as well. Further experiments to investigate the influence of HEAT on the activity of other coronaviruses need to be conducted to confirm the potential of HEAT as a pan-coronavirus inhibitor.

Currently most structurally well described SARS-CoV-2 M^pro inhibitors are rationally designed peptidomimetic or peptide-based inhibitors (*1*, *2*, *13*, *14*). However, some structures of non-peptidic inhibitors were recently reported. For example Baicalein is a natural product derived from traditional Chinese medicine (*15*). It is non-covalently bound in the active site of M^pro. In contrast the antineoplastic drug carmofur is converted by M^pro in an unknown manner and its fatty acid tail remains covalently bound to Cys145 (*16*). Chemically the mechanism of reaction of carmofur with M^pro is different from the one we observe for HEAT.

In conclusion we have discovered that the alpha adrenergic antagonist HEAT is metabolized by and covalently bound to the active site of SARS-CoV-2 M^pro, but does not exhibit an inhibitory effect on the enzyme activity. In spite of this, the complex between METT and M^pro can be considered a starting point for future drug design efforts. Chemical modifications of the molecule might help to endow it with pharmacologically relevant inhibitory activity and to develop HEAT towards a viable drug option.

The obtained results also strongly highlight that drug discovery investigations to treat SARS-Cov-2 should not only focus on traditional and conservative approaches aiming to identify substrate-analogue inhibitors, but should also include the identification and discovery of pro-drug compounds that could act as suicide drugs (*17*, *18*). This approach is based on the simple but effective idea of a classical pro-drug, which needs to be metabolized first before becoming toxic for the infecting agent, in this case the virus SARS Cov2 via deactivation of M^pro.

In this context and according to our postulated model, HEAT can be considered as a pro-drug as well as suicide drug which binds first to the active site of M^pro, where it is cleaved in an E1cB-like reaction mechanism into 2-methylene-1-tetralone and tyramine. The Cys145 sulfur subsequently binds covalently in a Michael addition to the methylene carbon atom of 2-methylene-1-tetralone thereby blocking the active site and inactivating M^pro.

Methods

M^pro crystallization. The protein was overexpressed in *E. coli* and purified for subsequent crystallization according to previously published methods (*1*). Co-crystallization with the compounds was achieved by equilibrating a 6.25 mg/ml protein solution in 20 mM HEPES buffer (pH 7.8) containing 1 mM DTT, 1 mM EDTA, and 150 mM NaCl against a reservoir solution of 100 mM MIB, pH 7.5, containing 25% w/w PEG 1500 and 5% *v/v* DMSO. Prior to crystallization compound solutions in DMSO were dried onto the wells of SwissCI 96-well plates. To obtain well-diffracting crystals in a reproducible way seeding was applied for crystal growth. Crystals appeared within a few hours and reached their final size after 2 - 3 days. Crystals were manually harvested and flash-frozen in liquid nitrogen for subsequent X-ray diffraction data collection.

X-ray data collection and processing. X-ray data were collected at beamline P11 at the synchrotron radiation facility PETRA III at DESY in Hamburg(*19*). Data were collected at an X-ray energy of 12 keV. Using the P11 sample changing robot data collection time was 3 minutes for a full 200 degree rotation dataset, including sample exchange. Data analysis was carried out using XDS (*20*), COOT (*21*), PHENIX (*22*), and PanDDA (*23*), in addition to some bespoke software tools. Chemical restraints for the bound compound were generated from SMILES using phenix.elbow (*24*) and compared with results from acedrg (*25*) and grade (*26*). The fully refined model and structure factors were deposited in the PDB with accession code 6YNQ. The distribution of protein residues over the Ramachandran plot was 97.04/2.63/0.33% (favored/allowed/outliers).

Mass-spectrometric analysis of compound binding. To prepare samples of M^{pro} for native MS measurements, they were buffer-exchanged into ESI compatible solutions. M^{pro} was buffer exchanged into 300 mM NH_4OAc, pH 7.5 by five cycles of centrifugal gel filtration (Vivaspin columns, 30.000 MWCO, Sartorius). M^{pro} (250 µM, 300 mM NH_4OAc, pH 7.5) was incubated with HEAT (2.5 mM in DMSO) for 20 h and diluted 1:20.
Nano ESI capillaries were pulled in-house from borosilicate capillaries (1.2 mm outer diameter, 0.68 mm inner diameter, filament, World Precision Instruments) with a micropipette puller (P-1000, Sutter instruments) using a squared box filament (2.5 × 2.5 mm^2, Sutter Instruments) in a two-step program. Subsequently, capillaries were gold-coated using a sputter coater (Q150R, Quorum Technologies) with 40 mA, 200 s, tooling factor 2.3 and end bleed vacuum of $8×10^{-2}$ mbar argon. Native MS was performed using an electrospray quadrupole time-of-flight (ESI-Q-TOF) instrument (Q-TOF2, Micromass/Waters, MS Vision) modified for higher masses (*27*). Samples were ionized in positive ion mode with voltages applied at the capillary of 1300 V and at the cone of 130 V. The pressure in the source region was kept at 10 mbar throughout all native MS experiments. For desolvation and dissociation, the pressure in the collision cell was adjusted to $1.8×10^{-2}$ mbar argon. Native-like spectra were obtained at an accelerating voltage of 30-100 V, while for collision-induced dissociation these voltages were increased to 150-170 V. In ESI-MS overview spectra, the quadrupole profile was 1-10,000 m/z. In tandem MS, for precursor selection, LMres and HMres were adjusted at 50 V collisional voltage until a single peak was recorded, then dissociation was induced. To calibrate raw data, CsI (25 mg/ml) spectra were acquired and calibration was carried out with MassLynx (Waters) software. Data were analysed using MassLynx (Waters).

Small-molecule mass spectrometry. Analytical LC-MS data were obtained using a Waters Alliance 2795-HT LC-MS system equipped with a Phenomenex Kinetex column (2.6 µm, C_{18}, 100 Å, 4.6 x 100 mm^2) and a Phenomenex Security Guard precolumn (Luna, C5, 300 Å) at a flow rate of 1 mL/min at 30 ˚C. Detection was carried out by a diode array detector (Waters 2998) in the range 210 to 600 nm together with a Waters Quatro-Micro mass detector, operating simultaneously in ES+ and ES− modes between 100 and 1000 m/z. Solvents were: A, HPLC grade H_2O containing 0.05% formic acid;

B, HPLC grade CH_3CN containing 0.045% formic acid. The gradient was as follows: 0 min, 10% B; 10 min, 90% B; 12 min, 90% B; 13 min, 10% B; 15 min, 10% B. Samples (20 µl) were made up in a mixture of buffer and CH_3CN between 10 and 1000 µg/mL. Enzyme assays were terminated by addition of an equivalent volume of CH_3CN. Protein was removed by centrifugation and reaction components were directly injected.

References:
1. L. Zhang, D. Lin, X. Sun, U. Curth, C. Drosten, L. Sauerhering, S. Becker, K. Rox, R. Hilgenfeld, Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved α-ketoamide inhibitors. *Science*. **368**, 409–412 (2020).

2. Z. Jin, X. Du, Y. Xu, Y. Deng, M. Liu, Y. Zhao, B. Zhang, X. Li, L. Zhang, C. Peng, Y. Duan, J. Yu, L. Wang, K. Yang, F. Liu, R. Jiang, X. Yang, T. You, X. Liu, X. Yang, F. Bai, H. Liu, X. Liu, L. W. Guddat, W. Xu, G. Xiao, C. Qin, Z. Shi, H. Jiang, Z. Rao, H. Yang, Structure of M pro from COVID-19 virus and discovery of its inhibitors. *Nature*, 1–9 (2020).

3. R. Hilgenfeld, From SARS to MERS: crystallographic studies on coronaviral proteases enable antiviral drug design. *FEBS J.* **281**, 4085–4096 (2014).

4. Identification of inhibitors of SARS-CoV-2 in-vitro cellular toxicity in human (Caco-2) cells using a large scale drug repurposing collection (2020), doi:10.21203/rs.3.rs-23951/v1.

5. S. M. Corsello, J. A. Bittker, Z. Liu, J. Gould, P. McCarren, J. E. Hirschman, S. E. Johnston, A. Vrcic, B. Wong, M. Khan, J. Asiedu, R. Narayan, C. C. Mader, A. Subramanian, T. R. Golub, The Drug Repurposing Hub: a next-generation drug library and information resource. *Nat. Med.* **23**, 405–408 (2017).

6. J. Schiebel, N. Radeva, S. G. Krimmer, X. Wang, M. Stieler, F. R. Ehrmann, K. Fu, A. Metz, F. U. Huschmann, M. S. Weiss, U. Mueller, A. Heine, G. Klebe, Six Biophysical Screening Methods Miss a Large Proportion of Crystallographically Discovered Fragment Hits: A Case Study. *ACS Chem. Biol.* **11**, 1693–1701 (2016).

7. DE1913199C3, (available at http://depatisnet.dpma.de/DepatisNet/depatisnet?action=bibdat&docid=DE00000 1913199C3).

8. H. Glossmann, F. Lübbecke, B. Habermann, [125I]-HEAT: Fifty percent of the ligand can bind to the alpha1-adrenoceptors with extremely high affinity. *Naunyn. Schmiedebergs Arch. Pharmacol.* **321**, 7–10 (1982).

9. R. M. DeMarinis, M. Wise, J. P. Hieble, R. R. Ruffolo, in *The alpha-1 Adrenergic Receptors*, R. R. Ruffolo, Ed. (Humana Press, Totowa, NJ, 1987; https://doi.org/10.1007/978-1-4612-4582-7_6), *The Receptors*, pp. 211–265.

10. J. Dülfer, A. Kadek, J.-D. Kopicki, B. Krichel, C. Uetrecht, in *Advances in Virus Research*, F. A. Rey, Ed. (Academic Press, 2019; http://www.sciencedirect.com/science/article/pii/S006535271930020X), vol. 105 of *Complementary Strategies to Understand Virus Structure and Function*, pp. 189–238.

11. N. Schneider, G. Lange, S. Hindle, R. Klein, M. Rarey, A consistent description of HYdrogen bond and DEhydration energies in protein–ligand complexes: methods behind the HYDE scoring function. *J. Comput. Aided Mol. Des.* **27**, 15–29 (2013).

12. J. Ziebuhr, E. J. Snijder, A. E. Gorbalenya, Virus-encoded proteinases and proteolytic processing in the Nidovirales. *J. Gen. Virol.* **81**, 853–879 (2000).

13. A. D. Mescar, A taxonomically-driven approach to development of potent, broad-spectrum inhibitors of coronavirus main protease including SARS-CoV-2 (COVID-19). *Be Publ.*

14. W. Dai, B. Zhang, H. Su, J. Li, Y. Zhao, X. Xie, Z. Jin, F. Liu, C. Li, Y. Li, F. Bai, H. Wang, X. Cheng, X. Cen, S. Hu, X. Yang, J. Wang, X. Liu, G. Xiao, H. Jiang, Z. Rao, L.-K. Zhang, Y. Xu, H. Yang, H. Liu, Structure-based design of antiviral drug candidates targeting the SARS-CoV-2 main protease. *Science* (2020), doi:10.1126/science.abb4489.

15. H. Su, S. Yao, W. Zhao, M. Li, J. Liu, W. Shang, H. Xie, C. Ke, M. Gao, K. Yu, H. Liu, J. Shen, W. Tang, L. Zhang, J. Zuo, H. Jiang, F. Bai, Y. Wu, Y. Ye, Y. Xu, *bioRxiv*, in press, doi:10.1101/2020.04.13.038687.

16. Z. Jin, Y. Zhao, Y. Sun, B. Zhang, H. Wang, Y. Wu, Y. Zhu, C. Zhu, T. Hu, X. Du, Y. Duan, J. Yu, X. Yang, X. Yang, X. Liu, L. W. Guddat, G. Xiao, L. Zhang, H. Yang, Z. Rao, *bioRxiv*, in press, doi:10.1101/2020.04.09.033233.

17. J. Drebes, M. Künz, B. Windshügel, A. G. Kikhney, I. B. Müller, R. J. Eberle, D. Oberthür, H. Cang, D. I. Svergun, M. Perbandt, C. Betzel, C. Wrenger, Structure of ThiM from Vitamin B1 biosynthetic pathway of Staphylococcus aureus – Insights into a novel pro-drug approach addressing MRSA infections. *Sci. Rep.* **6**, 1–10 (2016).

18. J. Drebes, M. Kunz, C. A. Pereira, C. B. and C. Wrenger, MRSA Infections: From Classical Treatment to Suicide Drugs. *Curr. Med. Chem.* (2014), (available at http://www.eurekaselect.com/118131/article).

19. A. Burkhardt, T. Pakendorf, B. Reime, J. Meyer, P. Fischer, N. Stübe, S. Panneerselvam, O. Lorbeer, K. Stachnik, M. Warmer, P. Rödig, D. Göries, A. Meents, Status of the crystallography beamlines at PETRA III. *Eur. Phys. J. Plus*. **131**, 56 (2016).

20. W. Kabsch, XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132 (2010).

21. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of Coot. *Acta Crystallogr Biol Crystallogr*. **66**, 486–501 (2010).

22. P. D. Adams, P. V. Afonine, G. Bunkóczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L.-W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, P. H. Zwart, PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).

23. N. M. Pearce, T. Krojer, A. R. Bradley, P. Collins, R. P. Nowak, R. Talon, B. D. Marsden, S. Kelm, J. Shi, C. M. Deane, F. von Delft, A multi-crystal method for extracting obscured crystallographic states from conventionally uninterpretable electron density. *Nat. Commun.* **8**, 1–8 (2017).

24. N. W. Moriarty, R. W. Grosse-Kunstleve, P. D. Adams, electronic Ligand Builder and Optimization Workbench (eLBOW): a tool for ligand coordinate and restraint generation. *Acta Crystallogr. D Biol. Crystallogr.* **65**, 1074–1080 (2009).

25. F. Long, R. A. Nicholls, P. Emsley, S. Gražulis, A. Merkys, A. Vaitkus, G. N. Murshudov, AceDRG: a stereochemical description generator for ligands. *Acta Crystallogr. Sect. Struct. Biol.* **73**, 112–122 (2017).

26. Grade Web Server, (available at http://grade.globalphasing.org/cgi-bin/grade/server.cgi).

27. R. H. H. van den Heuvel, E. van Duijn, H. Mazon, S. A. Synowsky, K. Lorenzen, C. Versluis, S. J. J. Brouns, D. Langridge, J. van der Oost, J. Hoyes, A. J. R. Heck, Improving the Performance of a Quadrupole Time-of-Flight Instrument for Macromolecular Mass Spectrometry. *Anal. Chem.* **78**, 7473–7483 (2006).

Author list:

Sebastian Guenther[1,*,#], Patrick Y. A. Reinke[1,*], Dominik Oberthuer[1], Oleksandr Yefanov[1], Helen Ginn[2], Susanne Meier[3,4], Thomas J. Lane[5], Kristina Lorenzen[6], Luca Gelisio[1], Wolfgang Brehm[1], Illona Dunkel[7], Martin Domaracky[1], Sofiane Saouane[8], Julia Lieske[1], Christiane Ehrt[9], Faisal Koua[1], Alexandra Tolstikova[1], Thomas White[1], Michael Groessler[1], Holger Fleckenstein[1], Fabian Trost[1], Marina Galchenkova[1], Yaroslav Gevorkov[1], Chufeng Li[1], Salah Awel[1], Ariana Peck[10], Paulraj Lourdu Xavier[1], Miriam Barthelmess[1], Frank Schlünzen[1], Nadine Werner[11], Hina Andaleeb[11], Najeeb Ullah[11], Sven Falke[11], Bruno Alves Franca[11], Martin Schwinzer[11], Hévila Brognaro[11], Brandon Seychell[12], Henry Gieseler[3,4], Diogo Melo[3,4], Jo J. Zaitsev-Doyle[3,4], Brenna Norton-Baker[13], Juraj Knoska[1], Gisel Esperanza[1], Aida Rahmani Mashhour[1], Filip Guicking[1], Vincent Hennicke[1], Pontus Fischer[1], Cromarte Rogers[3,4], Diana C. F. Monteiro[14], Johanna Hakanpää[8], Jan Meyer[8], Heshmat Noei[8], Phil Gribbon[15], Bernhard Ellinger[15], Maria Kuzikov[15], Markus Wolf[15], Linlin

Zhang[16,17], Xinyuanyuan Sun[16,17], Jonathan Pletzer-Zelgert[9], Jan Wollenhaupt[18], Christian Feiler[18], Manfred Weiss[18], Eike-Christian Schulz[13], Pedram Mehrabi[13], Christina Schmidt[6], Robin Schubert[6], Huijong Han[6], Boris Krichel[19], Yaiza Fernandez Garcia[20], Beatriz Escudero-Pérez[20], Stephan Günther[20], Dusan Turk[21], Charlotte Uetrecht[19], Tobias Beck[3,12], Henning Tidow[3,12], Ashwin Chari[22], Andrea Zaliani[15], Matthias Rarey[19], Russell Cox[23], Rolf Hilgenfeld[17,24], Henry N. Chapman[1,3], Arwen R. Pearson[3,4], Christian Betzel[3,11], and Alke Meents[1,#]

(1) Center for Free-Electron Laser Science, DESY, Notkestrasse 85, 22607 Hamburg, Germany. alke.meents@desy.de

(2) Diamond Light Source Ltd. Diamond House, Harwell Science and Innovation Campus, Didcot, OX11 0DE, UK.

(3) Hamburg Centre for Ultrafast Imaging, Universität Hamburg, Luruper Chaussee 149, 22761 Hamburg, Germany

(4) Universität Hamburg, Institut für Nanostruktur- und Festkörperphysik, Luruper Chaussee 149, 22761 Hamburg, Germany

(5) Linac Coherent Light Source, SLAC National Accelerator Laboratory, Menlo Park, California, USA

(6) European XFEL GmbH. Holzkoppel 4, 22869 Schenefeld, Germany

(7) Max Planck Institute for Molecular Genetics, Ihnestraße 63-73, 14195 Berlin, Germany

(8) Deutsches Elektronen Synchrotron (DESY), Photon Science, Notkestrasse 85, 22607, Hamburg, Germany.

(9) Universität Hamburg, Center for Bioinformatics, Bundesstr. 43, 20146 Hamburg, Germany

(10) Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA 91125, USA

(11) Universität Hamburg, Department of Chemistry, Institute of Biochemistry and Molecular Biology and Laboratory for Structural Biology of Infection and Inflammation, c/o DESY, 22607 Hamburg, Germany

(12) Universität Hamburg, Department of Chemistry, Institute of Biochemistry and Molecular Biology, Martin-Luther-King-Platz 6, 20146 Hamburg, Germany

(13) Max Planck Institute for the Structure and Dynamics of Matter, Luruper Chaussee 149, 22761 Hamburg, Germany

(14) Hauptmann Woodward Medical Research Institute, 700 Ellicott Street, Buffalo, NY, 14203, USA

(15) Fraunhofer Institute for Molecular Biology and Applied Ecology (IME), screening port, Schnackenburgallee 114, 22525 Hamburg, Germany

(16) Institute of Biochemistry, Center for Structural and Cell Biology in Medicine, University of Lübeck, 23562 Lübeck, Germany.

(17) German Center for Infection Research (DZIF), Hamburg-Lübeck-Borstel-Riems Site, University of Lübeck, 23562 Lübeck, Germany

(18) Helmholtz Zentrum Berlin, Hahn-Meitner-Platz 1, 14109 Berlin, Germany

(19) Dynamics of Viral Structures, Heinrich-Pette-Institut, Leibniz-Institut für Experimentelle Virologie, Martinistraße 52, 20251 Hamburg, Germany

(20) Bernhard Nocht Institute for Tropical Medicine, Bernhard-Nocht-Straße 74, 20359 Hamburg, Germany

(21) Department of Biochemistry & Molecular & Structural Biology, Jozef Stefan Institute, Jamova 39, 1 000 Ljubljana, Slovenia

(22) Department of Structural Dynamics, Max Planck Institute for Biophysical Chemistry, Am Fassberg 11, 37077 Göttingen, Germany

(23) Institute for Organic Chemistry and BMWZ, Leibniz University of Hannover, Schneiderberg 38, 30167 Hannover, Germany

(24) Laboratory for Antiviral Chemotherapy, Institute of Chemistry & Metabolomics and Institute of Molecular Medicine, University of Lübeck, Ratzeburger Allee 160, 23562 Lübeck, Germany
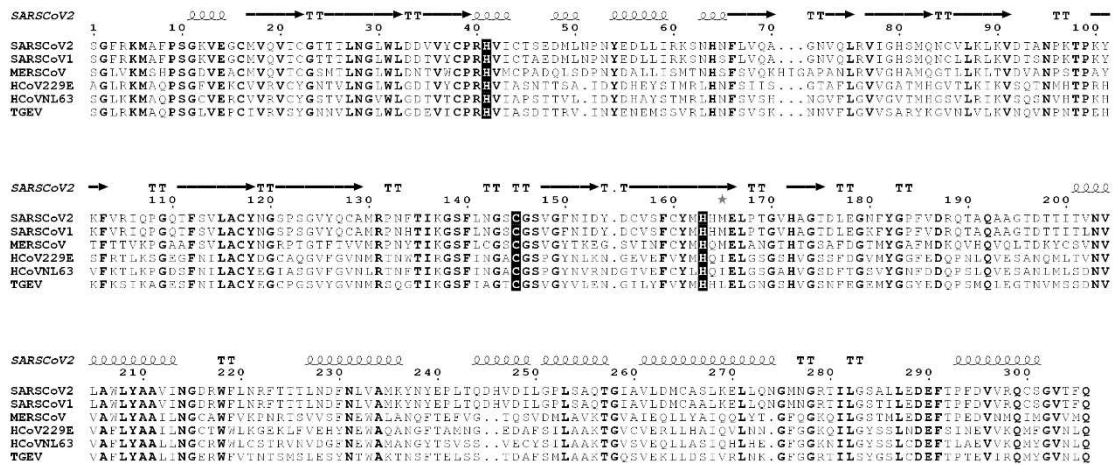
*) These authors contributed equally.
#) corresponding authors

Table 1 Data collection and refinement statistics (molecular replacement)

	M^{pro}/2-methyl-1-tetralone
PDB	6YNQ
Data collection	
Space group	C2
Cell dimensions	
a, b, c (Å)	113.08, 53.15, 44.65
α, β, γ(°)	90, 102.96, 90
Resolution (Å)	27.29 - 1.8 (1.864 - 1.8)
R_{sym} or R_{merge}	0.113 (1.853)
CC1/2	0.997 (0.435)
I / σI	8.1 (0.9)
Completeness (%)	97.37 (95.85)
Redundancy	3.8 (3.9)
Refinement	
Resolution (Å)	27.29 - 1.8
No. reflections	23506 (2309)
R_{work} / R_{free}	0.1821/0.2285
No. of non-hyrodgen atoms	2839
Protein	2528
Ligand/ion	17
Solvent	294
B-factors	
Protein	33.01
Ligand/ion	44.01
Solvent	38.3
R.m.s. deviations	
Bond lengths (Å)	0.011
Bond angles (°)	0.97

*Values in parentheses are for highest-resolution shell.
Data was collected from a single crystal

Supplementary figure 1: *Conservation of M^{pro} active site in α- (HCoV229E, HCovNL63, TGEV) and β- (SARS-CoV-1&2, MERS-CoV) coronaviruses.* Conserved residues are depicted in bold. Active site residues His41 and Cys145 are highlighted in black box. Also His163 forming a hydrogen bond to METT is conserved.

Supplementary figure 2:

To be added