bioRxiv preprint doi: https://doi.org/10.1101/2020.11.12.378422; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Massive X-ray screening reveals two allosteric drug binding sites of SARS-CoV-2 main protease

Sebastian Günther(1,*,#), Patrick Y. A. Reinke(1,*), Yaiza Fernández-García(2), Julia Lieske(1), Thomas J. Lane(1), Helen Ginn(3),Faisal H. M. Koua(1), Christiane Ehrt(4), Wiebke Ewert(1), Dominik Oberthuer(1), Oleksandr Yefanov(1), Susanne Meier(5,6), Kristina Lorenzen(7), Boris Krichel(8), Janine-Denise Kopicki(8), Luca Gelisio(1), Wolfgang Brehm(1), Ilona Dunkel(9), Brandon Seychell(10), Henry Gieseler(5,6), Brenna Norton-Baker(11), Beatriz Escudero-Pérez(2), Martin Domaracky(1), Sofiane Saouane(12), Alexandra Tolstikova(1), Thomas A. White(1), Anna Hänle(1), Michael Groessler(1), Holger Fleckenstein(1), Fabian Trost(1), Marina Galchenkova(1), Yaroslav Gevorkov(1,13), Chufeng Li(1), Salah Awel(1), Ariana Peck(14), Miriam Barthelmess(1), Frank Schlünzen(1),P. Lourdu Xavier(1,11), Nadine Werner(15), Hina Andaleeb(15), Najeeb Ullah(15), Sven Falke(15), Vasundara Srinivasan(15) Bruno Alves Franca(15), Martin Schwinzer(15), Hévila Brognaro(15), Cromarte Rogers(5,6), Diogo Melo(5,6), Jo J. Zaitsev-Doyle(5,6), Juraj Knoska(1), Gisel E. Peña Murillo(1), Aida Rahmani Mashhour(1), Filip Guicking(1), Vincent Hennicke(1), Pontus Fischer(1), Johanna Hakanpää(12), Jan Meyer(12), Phil Gribbon(16), Bernhard Ellinger(16), Maria Kuzikov(16), Markus Wolf(16), Gleb Borenkov(17), David von Stetten(17), Guillaume Pompidor(17), Isabel Bento(17), Saravanan Panneerselvam(17), Ivars Karpics(17), Thomas R. Schneider(17), Maria Marta Garcia Alai(17), Stephan Niebling(17), Christian Günther(17), Christina Schmidt(7), Robin Schubert(7), Huijong Han(7), Juliane Boger(18) Diana C. F. Monteiro(19), Linlin Zhang(18,20), Xinyuanyuan Sun(18,20), Jonathan Pletzer-Zelgert(4), Jan Wollenhaupt(21), Christian G. Feiler(21), Manfred S. Weiss(21), Eike-Christian Schulz(11), Pedram Mehrabi(11), Katarina Karničar (22,23), Aleksandra Usenik (22,23), Jure Loboda (22), Henning Tidow(5,24), Ashwin Chari(25), Rolf Hilgenfeld(18,20), Charlotte Uetrecht(8), Russell Cox(26), Andrea Zaliani(16), Tobias Beck(5,10), Matthias Rarey(4), Stephan Günther(2), Dusan Turk(22,23), Winfried Hinrichs (15), Henry N. Chapman(1,5), Arwen R. Pearson(5,6), Christian Betzel(5,15), and Alke Meents(1,#)

(1) Center for Free-Electron Laser Science, DESY, Notkestrasse 85, 22607 Hamburg, Germany.
(2) Bernhard Nocht Institute for Tropical Medicine, Bernhard-Nocht-Straße 74, 20359 Hamburg, Germany
(3) Diamond Light Source Ltd. Diamond House, Harwell Science and Innovation Campus, Didcot, OX11 0DE, UK.
(4) Universität Hamburg, Center for Bioinformatics, Bundesstr. 43, 20146 Hamburg, Germany
(5) Hamburg Centre for Ultrafast Imaging (CUI), Universität Hamburg, Luruper Chaussee 149, 22761 Hamburg, Germany
(6) Universität Hamburg, Institut für Nanostruktur- und Festkörperphysik, Luruper Chaussee 149, 22761 Hamburg, Germany
(7) European XFEL GmbH. Holzkoppel 4, 22869 Schenefeld, Germany
(8) Heinrich Pette Institute, Leibniz Institute for Experimental Virology, Martinistraße 52, 20251 Hamburg, Germany
(9) Max Planck Institute for Molecular Genetics, Ihnestraße 63-73, 14195 Berlin, Germany

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.12.378422; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

(10) Universität Hamburg, Department of Chemistry, Institute of Physical Chemistry, Grindelallee 117, 20146 Hamburg, Germany
(11) Max Planck Institute for the Structure and Dynamics of Matter, Luruper Chaussee 149, 22761 Hamburg, Germany
(12) Deutsches Elektronen Synchrotron (DESY), Photon Science, Notkestrasse 85, 22607, Hamburg, Germany.
(13) Vision Systems, Hamburg University of Technology, 21071 Hamburg, Germany
(14) Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA 91125, USA
(15) 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
(16) Fraunhofer Institute for Translational Medicine and Pharmacology (ITMP), Schnackenburgallee 114, 22525 Hamburg, Germany
(17) EMBL Outstation Hamburg, c/o DESY, Notkestrasse 85, 22607 Hamburg, Germany
(18) Institute of Molecular Medicine, University of Lübeck, 23562 Lübeck, Germany.
(19) Hauptmann Woodward Medical Research Institute, 700 Ellicott Street, Buffalo, NY, 14203, USA
(20) German Center for Infection Research (DZIF), Hamburg-Lübeck-Borstel-Riems Site, University of Lübeck, 23562 Lübeck, Germany
(21) Helmholtz Zentrum Berlin, Macromolecular Crystallography, Albert-Einstein-Str. 15, 12489 Berlin, Germany
(22) Department of Biochemistry & Molecular & Structural Biology, Jozef Stefan Institute, Jamova 39, 1 000 Ljubljana, Slovenia
(23) Centre of excellence for Integrated Approaches in Chemistry and Biology of Proteins (CIPKEBIP), Jamova 39, 1 000 Ljubljana,Slovenia
(24) Universität Hamburg, Department of Chemistry, Institute of Biochemistry and Molecular Biology, Martin-Luther-King-Platz 6, 20146 Hamburg, Germany
(25) Research Group for Structural Bichemistry and Mechanisms, Department of Structural Dynamics, Max Planck Institute for Biophysical Chemistry, Am Fassberg 11, 37077 Göttingen, Germany
(26) Institute for Organic Chemistry and BMWZ, Leibniz University of Hannover, Schneiderberg 38, 30167 Hannover, Germany

*) These authors contributed equally.
#) corresponding authors (sebastian.guenther@desy.de, alke.meents@desy.de)

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.12.378422; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Summary

The coronavirus disease (COVID-19) caused by SARS-CoV-2 is creating tremendous health problems and economical challenges for mankind. To date, no effective drug is available to directly treat the disease and prevent virus spreading. In a search for a drug against COVID-19, we have performed a massive X-ray crystallographic screen of repurposing drug libraries[1] containing 5953 individual compounds against the SARS-CoV-2 main protease (M^{pro}), which is a potent drug target as it is essential for the virus replication[2]. In contrast to commonly applied X-ray fragment screening experiments with molecules of low complexity, our screen tested already approved drugs and drugs in clinical trials. From the three-dimensional protein structures, we identified 37 compounds binding to M^{pro}. In subsequent cell-based viral reduction assays, one peptidomimetic and five non-peptidic compounds showed antiviral activity at non-toxic concentrations. Interestingly, two compounds bind outside the active site to the native dimer interface in close proximity to the S1 binding pocket. Another compound binds in a cleft between the catalytic and dimerization domain of M^{pro}. Neither binding site is related to the enzymatic active site and both represent attractive targets for drug development against SARS-CoV-2. This X-ray screening approach thus has the potential to help deliver an approved drug on an accelerated time-scale for this and future pandemics[3].

Introduction

Infection of host cells by SARS-CoV-2 critically depends on the complex interplay of several molecular factors of both, the host and the virus[4,5]. Coronaviruses are RNA-viruses with a genome of approximately 30,000 nucleotides. The viral open-reading frames, essential for replication of the virus, are expressed as two overlapping, large polyproteins, which must be separated into functional subunits for replication and transcription activity[4]. This proteolytic cleavage, which is vital for viral reproduction, is primarily accomplished by the main protease (M^{pro}), also known as 3C-like protease $3CL^{pro}$ or nsp5. M^{pro} cleaves the viral polyprotein pp1ab at eleven distinct sites. The core cleavage motif is Leu-Gln⇓(Ser/Ala/Gly)[4]. M^{pro} possesses a chymotrypsin-like fold appended with a C-terminal helical domain, and harbors a catalytic dyad comprised of Cys145 and His41[4]. 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 and dimerization of the enzyme, with a dissociation constant of ~2.5 μM[4]. Due to its central and vital involvement in virus replication, M^{pro} is recognized as a prime target for antiviral drug discovery and compound screening activities aiming to identify and optimize drugs which can tackle coronavirus infections[6]. Indeed, a number of recent publications confirm the potential of targeting M^{pro} for inhibition of virus replication[4,5].

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.12.378422; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

A rational approach to the identification of new drugs is structure-based drug design[7,8]. The first step is target selection followed by biochemical and biophysical characterization of the target protein and, most importantly, its structure determination by X-ray crystallography, NMR or cryo-electron microscopy. This knowledge forms the basis for subsequent *in silico* screening of up to millions of potential drug molecules, leading to the identification of potentially binding compounds. The most promising candidates are then subjected to screening *in vitro* for biological activity. Lead structures are derived from common structural features of these biologically active compounds. Further chemical modifications of lead structures can then create a drug candidate that can be tested in animal models and, finally, clinical trials.

The identification of drug-binding sites on a pharmacologically relevant target protein and the identification of lead structures is often supported by fragment-screening, monitoring the interaction of small molecules with the target through various biophysical approaches including NMR and X-ray crystallography. These screens typically encompass a few hundred fragments of low chemical complexity[9].

In contrast to such fragment-screening experiments, screening of libraries containing several thousands of larger and more complex compounds are typically conducted through biochemical or biophysical assays which are regarded as more amenable for high-throughput measurements, such as fluorescence- or cell-based assays. Until now, methods yielding structural information about the compound-target complex were only applied to a small subset of previously identified “hits” due to the large effort required for such experiments.

Over the past few years, X-ray crystallography has matured into a fast and highly automated method[10,11]. Once crystallization conditions for a target are established, the screening experiment becomes straight-forward. With the availability of highly automated beamlines at latest generation synchrotron sources, the recent advances in detector technology, and well-established data processing methods, X-ray structure determination now typically takes only a few minutes, enabling screening of several hundred samples per day. With these developments, screening of entire libraries containing several thousand compounds by X-ray crystallography is now feasible.

Here we report on our massive X-ray crystallographic screen of SARS-CoV-2 M^{pro} against two repurposing libraries containing in total 5953 unique compounds from the “Fraunhofer IME Repurposing Collection”[1] , which is based on the BROAD institute repurposing library[12], and the “Safe-in-man” library from Dompé Farmaceutici S.p.A. Analysis of the derived electron-density maps showed 37 structures with bound compounds. Further validation by native mass spectrometry and viral reduction assays led to the identification of six compounds showing significant *in vitro* antiviral activity against SARS-CoV-2, including inhibitors binding at allosteric sites. Our results illustrate the power of this approach, provide new insights, and pave the way

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.12.378422; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

for new strategies to develop drugs that are active against the virus. In addition, now that this accelerated process is established, it can readily be applied to facilitate responses to future epidemics.

High-throughput X-ray screen

For the screening experiments of the two repurposing libraries against M^{pro}, the protein was over-expressed and purified as previously described[4]. In contrast to crystallographic fragment-screening experiments that use small molecules of low molecular weight typically below 200 Da, the repurposing libraries are chemically more complex and with compounds twice the molecular weight (Figure 1A) and thus likely to bind more specifically and with higher affinity[13]. Due to the higher molecular weights, we performed co-crystallization experiments instead of compound soaking into native crystals[14]. Each compound was co-crystallized with M^{pro} by adding the compounds to the crystallization plates prior to crystallization solutions. In order to obtain homogeneously sized and high-quality crystals, seeding was used. Crystals were grown at a physiological pH-value of 7.5 and typically appeared after 2-3 days.

X-ray data collection was performed at cryogenic temperatures at beamlines P11, P13 and P14 at the PETRA III storage ring at DESY. In total, datasets from 6288 crystals were collected over a period of four weeks. From the 5953 unique compounds in our screen, we obtained crystals in 3955 cases, out of these, 3228 yielded high-quality diffraction data to a resolution better than 2.5 Å. 1196 datasets were suitable for subsequent automated structure refinement followed by cluster analysis[15] and pan dataset density analysis (PanDDA)[16]. In total, 43 compounds were found that bound to M^{pro}. Seven of these compounds had maleate as a counterion and in these structures maleate was found in the active site but not the compounds themselves, resulting in 37 unique binders. A summary of these together with additional experimental information, is provided in Suppl. Table 1 and 2. For these 37 compounds, the binding mode could be unambiguously determined for 29 molecules, including ten that could be classified as covalent binders. The majority of hits were found in the active site of the enzyme, which is defined by the binding pockets for the natural peptide substrate[17]. Six of 16 active-site binders covalently bind as thioethers to Cys145, one compound binds covalently as a thiohemiacetal to Cys145, one is coordinated through a zinc ion and eight bind non-covalently. The remaining 13 compounds bind outside the active site at various locations (Figure 1B).

In vitro antiviral activity

Out of the 43 hits from our X-ray screen, 39 compounds were tested for their antiviral activity against SARS-CoV-2 in cell assays. Ten compounds reduced viral RNA replication by at least two orders of magnitude in Vero E6 cells (Figure S1). Six of the ten compounds show favorable cytotoxicity profiles with selectivity indexes ($SI = CC_{50} / EC_{50}$) greater than five. For the remaining four compounds, the antiviral activity cannot be unambiguously attributed to M^{pro} inhibition. The six active compounds (Figure 2) interact with M^{pro} at two other binding sites. Three compounds bind

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.12.378422; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

covalently to the active site and one non-covalently. Two compounds bind at the dimer interface between the two monomers. Another compound binds in a cleft between the catalytic and dimerization domain, but shows slightly weaker antiviral activity.

Active-site binding compounds

For three compounds **Tolperisone**, HEAT and isofloxythepin, breakdown products are observed to be covalently bound in the active site. Of these HEAT and isofloxythepin show activity but unfavorable cytotoxicity, whereas Tolperisone is active (EC_{50} = 17.16 ± 1.76 µM) and shows no cytotoxicity at 100 µM (Figure 2).

Tolperisone and HEAT are β-ketoamines. For both compounds, a breakdown product of the parent drug is observed to covalently bind as Michael-acceptor to the thiol of Cys145. Similarly, the aromatic ring system of both tolperisone (Figure 3A) and HEAT (Figure 3B) protrudes into the S1 pocket and forms van der Waals contacts with the backbone of Phe140 and Leu141 and the side chain of Glu166. In addition, the keto group accepts a hydrogen bond from the imidazole side chain of His163. Tolperisone and HEAT bind exclusively in the (S)-configuration. Interestingly, we only observe the part of the drug containing the activating ketone, while the remaining part with the amine group is missing in the electron-density maps. For HEAT, this binding mode has been confirmed independently by mass spectrometry (Figure S3 and Table S3). A similar observation has been reported for binding of β-ketoamines to type-1 methionine aminopeptidases, where the parent compound decomposes into an amine and an α,β-unsaturated ketone which subsequently binds to the thiol of the catalytic cysteine[18]. This is a typical situation for a pro-drug[19]. Tolperisone is in use as a skeletal muscle relaxant[20].

Isofloxythepin appears to similarly bind as a fragment to Cys145 (Figure 3C). Here, the piperazine group is not found in the crystal structure but the dibenzothiepine moiety is observed in the active site, bound as a thioether to Cys145. The tricyclic system stretches from the S1 across to the S1' pocket. According to the electron-density maps, two orientations of the molecule are possible, with either the fluorine or the isopropyl group placed inside the S1 pocket. Degradation of the drug with piperazine as the leaving group has been previously reported[21] and was confirmed by mass spectrometry (Figure S3). Isofloxythepin is an antagonist of dopamine receptors D1 and D2[22] and has been tested as a neuroleptic in phase II clinical trials.

Triglycidyl isocyanurate shows antiviral activity and adopts two binding modes to the M^{pro} active site, one covalent and one non-covalent. In both modes, the compound's central ring sits on top of the catalytic dyad (His41, Cys145) and its three epoxypropyl substituents reach into subsites S1', S1 and S2. The non-covalent binding mode is stabilized by hydrogen bonds to the main chain of Gly143 and Gly166, and to the side chain of His163. In the covalently bound form, one oxirane ring is opened by nucleophilic attack of Cys145 forming a thioether (Figure 3D). The

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.12.378422; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

use of epoxides as warheads for inhibition of M^{pro} offers another avenue for covalent inhibitors, whereas epoxysuccinyl warheads have been extensively used in biochemistry, cell biology and later in clinical studies[23]. Triglycidyl isocyanurate (teroxirone, Henkel's agent) has been tested as antitumor agent[24].

Maleate as counter ion of some of the applied compounds is observed as a Michael adduct in seven structures. These compounds exhibit no antiviral activity. A similar Michael adduct has been described for maleate isomerase[25] as an intermediate structure in the isomerization reaction (Figure S2B). A detailed description is given in the supplement.

Calpeptin shows the highest antiviral activity in the screen, with an EC_{50} value in the lower µM range. It binds covalently via its aldehyde group to Cys145, forming a thiohemiacetal. This peptidomimetic inhibitor occupies substrate pockets S1 to S3, highly similar to inhibitor GC-376[26,27], calpain inhibitors[28] and other peptidomimetic inhibitors such as N3[5] and the α-ketoamide 13b[4]. The peptidomimetic backbone forms hydrogen bonds to the main chain of His164 and Glu166, whereas the norleucine side chain is in van der Waals contacts with the backbone of Phe140, Leu141 and Asn142 (Figure 3E). Calpeptin has known activity against SARS-CoV-2[26]. The structure is highly similar to leupeptin, which served as positive control in our screen (Figure S2A). *In silico* docking experiments verified the peptidomimetic compound Calpeptin as a likely M^{pro} binding molecule (Table S4).

MUT056399 is the only active-site binding compound we found without a covalent bond to Cys145 but still reduced viral replication. The diphenyl ether core of MUT056399 blocks access to the catalytic site consisting of Cys145 and His41. The terminal carboxamide group occupies pocket S1 and forms hydrogen bonds to the side chain of His163 and the backbone of Phe140 (Figure 3F). The other part of the molecule reaches deep into pocket S2, which is enlarged by a shift of the side chain of Met49 out of the substrate binding pocket. MUT056399 was developed as an antibacterial agent against multidrug-resistant *Staphylococcus aureus* strains[29].

Two allosteric binding sites identified

In addition to the active site, as the most obvious target for drug development, we discovered two allosteric binding sites of M^{pro} which have previously not been reported. Five compounds of our X-ray screen, two of them showing antiviral activity in combination with low cytotoxicity (Figure 2) bind in a hydrophobic pocket in the C-terminal dimerization domain, located close to the oxyanion hole in pocket S2 of the substrate binding site. Another compound with slightly lower antiviral activity binds in between the catalytic and dimerization domains of M^{pro}.

Pelitinib, ifenprodil, RS-102895, PD-168568 and tofogliflozin bind to a hydrophobic pocket formed by helices of the C-terminal dimerization domain (Figure 4A).

Pelitinib shows the second highest antiviral activity in our screen, with an EC_{50} value of 1.25 µM. Its halogenated benzene ring to a hydrophobic groove in the helical

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.12.378422; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

dimerization domain, formed by Ile213, Leu253, Gln256, Val297 and Cys300 (Figure 4D). The central 3-cyanoquinoline moiety interacts with the end of the C-terminal helix (Ser301). The ethyl ether substituent pushes against Tyr118 and Asn142 (from loop 141-144 of the S1 pocket) of the opposing protomer within the native dimer. Pelitinib is known as an amine-catalyzed Michael acceptor[30], developed to bind to a cysteine in the active site of a tyrosine kinase. But from its observed binding position it is impossible for it to reach into the active site and no evidence for covalent binding to Cys145 is found in the electron-density maps. Pelitinib is an irreversible epidermal growth factor receptor inhibitor and developed as an anticancer agent[31].

Ifenprodil, RS-102895 and PD-168568 all exhibit an elongated structure, consisting of two aromatic ring systems separated by a linker containing a piperidine or piperazine ring (Fig 4B). All three compounds have a distance of at least 12 Å between the terminal aromatic rings. Thus, this binding mode is unlikely to be identified through fragment screening. The hydrophobic pocket in the helical domain is covered by the side chain of Gln256. In our complex structures, this side chain adopts a different conformation exposing Ile213, and generating the hydrophobic pocket. One of the terminal aromatic ring systems is inserted into the hydrophobic groove in the dimerization domain. The linker moiety stretches across the native dimer interface and the second aromatic ring is positioned close to Asn142, adjacent to the active site loop where residues 141-144 contribute to the pocket S1. In particular, in the case of RS-102895, two hydrogen bonds are formed to the side and main chains of Asn142. The exact interpretation of the binding mode in the crystal structures is complicated by the fact that the ligand is observed in two overlapping orientations created by the crystallographic twofold axis (Fig 4C). In contrast to ifenprodil, RS-102895 and PD-168568 do not exhibit selective antiviral activity (SI<5). All three compounds are GPCR antagonists. Ifenprodil antagonizes N-methyl-D-aspartate receptors[32], RS-102895 inhibits C-C chemokine receptor 2[33], and PD-168568 dopamine-receptors[34].

Tofogliflozin binds to the same hydrophobic pocket but no antiviral activity was observed at 100 µM, the highest concentration tested. In contrast to the previous four compounds, it does not reach across to the opposing protomer in the native dimer. Its main interaction with M^{pro} is via its isobenzofuran moiety that occupies the hydrophobic pocket (Figure S2S).

AT7519 is a unique hit in our screen as it binds in a deep groove between the catalytic domains and the dimerization domain (Figure 4E). The chlorinated benzene ring is engaged in various van der Waals interactions to loop 107-110, Val202, and Pro293 (Figure 4F). The central pyrazole has van der Waals contacts to Ile249, Phe294 and its adjacent carbonyl group forms a hydrogen bond to the side chain of Gln110. The terminal piperidine forms hydrogen bonds to the carboxylate of Asp153. This results in a displacement of loop 153-155, slightly narrowing the binding groove. The Cα-atom of Tyr154 moves by 2.8 Å, accompanied by a conformational change of Asp153. This allows hydrogen bonding to the compound and the formation of a salt-

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.12.378422; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

bridge to Arg298. In turn, Arg298 is crucial for dimerization[35]. The mutation Arg298Ala causes a reorientation of the dimerization domain relative to catalytic domain, leading to changes in the oxyanion hole and destabilization of the S1 pocket by the N-terminus. We assume that this binding site interferes with allosteric events required for enzymatic activity, for example substrate recognition. AT7519 was developed through fragment-based drug design against cyclin-dependent kinase 2 and was evaluated for treatment of human cancers[36] and shows weak antiviral activity but a poor selectivity index.

Further details about the remaining identified hit compounds are given in the supplement.

Discussion

Our X-ray screen revealed six compounds with previously unreported antiviral activity against SARS-CoV-2. Two of them, calpeptin and pelitinib, show strong antiviral activity combined with low cytotoxicity and are suitable for preclinical evaluation. The remaining compounds are valuable lead structures for further drug development.

The most active compound, calpeptin binds in the active site in the same way as other members of the large class of peptide-based inhibitors that bind as thiohemi-acetals or -ketals to M^{pro} [(2)]. However, in addition to this peptidomimetic inhibitor, we discovered several non-peptidic inhibitors. Those compounds binding to the active site of M^{pro} contained new Michael acceptors based on β-ketoamines (tolperisone and HEAT). These lead to the formation of thioethers and have not previously been described as prodrugs for viral proteases. We also identified a non-covalent binder, MUT056399, blocking the active site.

In addition, we discovered two allosteric drug binding sites outside the active site of M^{pro}. Pelitinib and four other hits bind at the hydrophobic pocket within the α-helical dimerization domain. According to our crystal structures, these binders extend out of the pocket and interact with loop 140-144. This loop is part of the S1 pocket and forms the oxyanion hole of the adjacent monomer within the native dimer. Previous work on M^{pro} of SARS-CoV demonstrated that the integrity of this pocket is crucial for enzyme activity[37].

Ifenprodil binds at the same allosteric site and shows antiviral activity, confirming this site is a suitable target for antivirals against SARS-CoV-2. Of note, ifenprodil is currently in phase IIb/III clinical trials for the treatment of COVID-19 based on the observation that it reduces mortality of lethal infection of H5N1 influenza in mice, likely through reduced inflammatory cytokine expression[38]. Our crystal structure and antiviral tests suggest an additional mode of action beyond this anti-inflammatory effect.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.12.378422; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

A comparison of coronavirus M^{pro} sequences shows that the compound binding residues of this allosteric site are conserved (Figure S4). Consequently, potential drugs targeting this allosteric binding site can be assumed to be robust against mutational variations and might also be effective against other coronaviruses.

The observed antiviral activity of AT7519, binding at the boundary of the dimerization and catalytic domain, demonstrates the potential of the second allosteric site as a druggable target.

Since the beginning of the pandemic, numerous screening campaigns using different approaches to target M^{pro}, including X-ray fragment screening and enzymatic activity, have been reported[39–42]. Remarkably, all non-peptidic inhibitors discovered in our massive X-ray screening effort have not previously been identified as active compounds. This highlights the benefit of using higher-molecular weight compounds with their potentially increased specificity and higher affinity to the biological target. A general advantage of using drug-repurposing libraries for such a screening is the proven bioactivity of the compounds and key properties such as cell-permeability are usually known[3].

The experimental methods and data analysis strategies of this massive X-ray screening of a protein against a repurposing library were built and optimized on the fly. Some similar campaigns were conducted at other facilities[39]. With the learning we and others gained we are now able to conduct such efforts in a streamlined fashion expending only a fraction of the resources initially needed. We now routinely measure 450 datasets per day on a single synchrotron beamline. Further improvements, such as advanced sample delivery techniques and employing artificial intelligence for data analysis, will allow us to further increase capacity towards our ultimate goal of collecting and analyzing data from more than 1000 samples per day. This approach will provide a fast-response platform prepared for future epidemics.

Methods

Protein production and purification

The protein was overexpressed in *E. coli* and purified for subsequent crystallization according to previously published protocols and plasmid constructs[4]. Lysis was carried out in 20 mM HEPES buffer supplemented with 150 mM NaCl using ultrasound for cell disruption. After separation of the cell fragments and the dissolved protein, a subsequent nickel NTA column was used to extract the M^{pro}-histidine-tag fusion. The cleavage of the histidine tag was achieved by a 3C protease during an overnight dialysis step. The histidine tag and the 3C protease were removed using a nickel NTA column, and as a final step a gel filtration was performed with an S200 Superdex column.

Crystallization experiments

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.12.378422; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Co-crystallization with the compounds was achieved mixing 0.23 µL of protein solution (6.25 mg/mL) in 20 mM HEPES buffer (pH 7.8) containing 1 mM DTT/TCEP (respectively), 1 mM EDTA, and 150 mM NaCl with 0.22 µL of reservoir solution consisting of 100 mM MIB, pH 7.5, containing 25% w/w PEG 1500 and 5% (v/v) DMSO, and 0.05 µL of a micro-seed crystal suspension. This growth solution was equilibrated by sitting drop vapor diffusion against 40 µL reservoir solution.

Prior to crystallization 125 nL droplets of 10 mM compound solutions from the two libraries in DMSO were applied to the wells of SwissCI 96-well plates (2-well or 3-well low profile, respectively) and subsequently dried in vacuum. Taking the crystallization drop volume into account this resulted in a final compound concentration of 2.5 mM and a molar ratio of ~13.6 of compound to protein. To obtain well-diffracting crystals in a reproducible way micro-seeding was applied for crystal growth[43]. Crystals appeared within a few hours and reached their final size (~200×100×10 µm^3) after 2 - 3 days. Crystals were manually harvested and flash-frozen in liquid nitrogen for subsequent X-ray diffraction data collection. We aimed at harvesting two crystals per crystallization condition as a compromise between through-put and increasing the probability to collect data from well diffracting crystals.

Data collection

Data collection was performed at beamlines P11, P13 and P14 at the PETRA III storage ring at DESY in Hamburg within a period of four weeks. Exclusive use of DESY beamline P11 was generously granted by the DESY directorate for the project.

Data processing and structure refinement

An automatic data processing and structure refinement pipeline “xia2pipe” as written specifically to support this project. Raw diffraction images from the PETRA III beamlines were processed using three crystallographic integration software packages: XDS[44], autoPROC[45] followed by staraniso[46], and DIALS via xia2[47,48]. Diffraction data quality indicators for all datasets and the 43 hits are summarized in Suppl. Figure S3. In total, 7857 unique crystals were harvested and frozen, of which 7258 were studied by X-ray diffraction at PETRA-III. Of these, 5934 produced diffraction data consistent with a protein lattice and were labeled as “successful” experiments. In some cases, multiple datasets were collected on a single crystal, so in total 8304 diffraction experiments were conducted with 6831 successful protein diffraction datasets obtained. As processed by DIALS, these 6831 datasets had an average resolution of 2.12 Å (criterion: CC1/2 > 0.5), CC1/2 of 0.97, and Wilson B of 27.8 Å^2 (Suppl. Figure S5). Crystallographic data of all structures submitted to the PDB are summarized in Suppl. Table S2.

For clustering and hit identification, all datasets were integrated and merged to a resolution of 1.7 Å. In order to reduce the influence of noise for lower resolution datasets, the following processing was applied to standardize the Wilson plot for each dataset: the datasets were split into equally sized bins, each covering 1000 reflections, and a linear fit was applied to the logarithm of the average intensities in

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.12.378422; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

each shell. The residual between the data and the Wilson fit was calculated, considering sequentially one additional bin from low to high resolution until the residual exceeded 10%, if applicable. The intensities in all higher resolution bins beyond this point were scaled to fit the calculated Wilson B factor.

The results of each dataset were then automatically refined using Phenix[49]. Refinement began by choosing one of two manually refined starting models (differing in their unit cell, Suppl. Table S2), selecting the starting model with the closest unit cell parameters, then proceeding in four steps: (1) rigid body and ADP refinement, (2) simulated annealing, ADP, and reciprocal space refinement, (3) real-space refinement, and (4) a final round of reciprocal space refinement as well as TLS refinement, with each residue pre-set as a TLS group. This procedure was hand-tuned on 5 test datasets; the procedure and parameters were manually adjusted to minimize Rfree until deemed satisfactory for the continuation of the project. All processing and refinement results were logged in a database, which enabled comparison between methods and improvement over time.

All code and parameters needed to reproduce this pipeline are available online[50].

Hitfinding: cluster4x and PanDDA Analysis

The resulting model structure Cα positions were then ingested into cluster4x[51], which briefly (a) computes a correlation coefficient between each structure over the position of all C_α atoms, (b) performs PCA the resulting correlation matrix, (c) presents 3 chosen principal components to a human, who then manually annotates clusters. Clusters were ordered chronologically and separated into groups of 1500 and subsequently clustered into groups of approximately 60-120 datasets based on a combination of reciprocal and Cα-atom differences using cluster4x. In an earlier version of the software, structure factor amplitudes were used for clustering instead of refined Cα positions, and both methods were applied for hitfinding. The resulting clusters were then analyzed via PanDDA[16] using default parameters. The resulting PanDDA analyses were manually inspected for hits which were recorded.

Manual structure refinement

Identified hits were further refined by alternating rounds of refinement using refmac[52], phenix.refine[49] or MAIN[53], interspersed with manual model building in COOT[54].

***In sillico* screening of compound libraries**

To enable a preselection of potentially promising compounds to support the experimental X-ray screening effort and to get an idea about the most promising compounds, we pursued a virtual screening workflow consisting of the selection of a representative ensemble of binding site conformations, non-covalent molecular docking and rescoring. We performed this study with 5,575 compounds of the Fraunhofer IME Repurposing Collection. UNICON[55] was applied to prepare the library compounds. To consider binding site flexibility, we used multiple receptor structures. We applied SIENA[56] to extract five representative binding site conformations for the active site of M^{pro}. We chose the structures with the PDB IDs 5RFH, 5RFO, 6W63, 6Y2G and 6YB7 The SIENA-derived aligned structures were

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.12.378422; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

used and the proteins were preprocessed using Protoss[57] to determine protonation states, tautomeric forms, and hydrogen orientations. The binding site was defined based on the active site ligand of the structure with the PDB ID 6Y2G (ligand ID O6K). A 12.5 Å radius of all ligand atoms was chosen as binding site definition. The new docking and scoring method JAMDA was applied with default settings for the five selected binding sites[58]. Subsequently, HYDE[59] was used for a rescoring of all predicted poses of the library compounds. The 200 highest ranked compounds of all 5,575 compounds according to the HYDE score were extracted. For 70 of these compounds, well-diffracting crystals were obtained in the X-ray screening. Intriguingly, only calpeptin, a known cysteine protease inhibitor, could be co-crystallized and was found on rank 3 (Table S4).

Mass Spectrometry

M^{pro} was prepared for native MS measurements by buffer-exchange into ESI compatible solutions (250 µM, 300 mM NH4OAc, 1 mM DTT, pH 7.5) by five cycles of centrifugal filtration (Vivaspin 500 columns, 30,000 MWCO, Sartorius). Inhibitors were dissolved to 1 mM in DMSO. Then Inhibitors and M^{pro} were mixed to final concentrations of 50 µM and 10 µM, respectively, and incubated for 16 h at 4 °C. For putative covalent ligands, compounds were incubated at 1 mM with 100 µM M^{pro} in 20 mM Tris, 150 mM NaCl, 1 mM TCEP, pH 7.8, for 16 h prior to buffer exchange. Buffer exchange was carried out as described above and samples were diluted tenfold prior to native MS measurements. All samples were prepared in triplicate. 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 (CCU-010, safematic) with 5.0 × 10-2 mbar, 30.0 mA, 100 s, 3 runs to vacuum limit 3.0 × 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[60]. Samples were ionized in positive ion mode with voltages of 1300 V applied at the capillary and of 130 V at the cone. 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 1.5 × 10^{-2} mbar argon. Native-like spectra were obtained at an accelerating voltage of 30 V. To calibrate raw data, CsI (25 mg/ml) spectra were acquired. Calibration and data analysis were carried out with MassLynx 4.1 (Waters) software. In order to determine each inhibitor binding to M^{pro}, peak intensities of zero, one or two bound ligands were analyzed from three independently recorded mass spectra at 30 V acceleration voltage. Results are shown in Supplementary Table S3.

Antiviral assays

Compounds. All compounds were diluted to a 50 mM concentration in 100% DMSO and stored at -80°C.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.12.378422; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Cytotoxicity assays. Vero E6 cells (ATCC CRL-1586) were seeded at 3.5×10^4 cells/well in 96-well plates. After 24 h, the cell culture media was changed and 2-fold serial dilutions of the compounds were added. Cell viability under 42 h compound treatment was determined via the Cell Counting Kit-8 (CCK-8, Sigma-Aldrich #96992) following the manufacturer´s instructions. The cytotoxic concentrations that reduced cell growth by 50% (CC_{50}) were calculated by fitting the data to the sigmoidal function using GraphPad Prism version 8.00 (GraphPad Software, La Jolla California USA, www.graphpad.com).

Antiviral activity assays. Vero E6 cells (ATCC CRL-1586) seeded at 3.5×10^4 cells/well in 96-well plates were pretreated 24 h later with twofold serial dilutions of the compounds. After 1 h incubation with the compounds, SARS-CoV-2 (strain SARS-CoV-2/human/DEU/HH-1/2020) was subsequently added at a MOI of 0.01 and allowed absorption for 1 h. The viral inoculum was removed, cells were washed with PBS without Mg^{2+} / Ca^{2+} and fresh media containing the compounds (final DMSO concentration 0.5% (v/v)) was added to the cells. Cell culture supernatant was harvest 42 hpi and stored at -80°C. Viral RNA was purified from the cell culture supernatant using the QIAamp Viral RNA Mini Kit (QIAGEN #52906) in accordance with the manufacturer´s instructions. Quantification of vRNA was carried out by the interpolation of RT-qPCR (RealStar SARS-CoV-2 RT-PCR Kit, Altona Diagnostics #821005) results onto a standard curve generated with serial dilutions of a template of known concentration. Titers of infectious virus particles were measured via immunofocus assay. Briefly, Vero E6 cells (ATCC CRL-1586) seeded at 3.5×10^4 cells/well in 96-well plates were inoculated with 50 µl of serial tenfold dilutions of cell culture supernatant from treated cells. The inoculum was removed after 1 h and replaced by a 1.5% methylcellulose-DMEM-5% FBS overlay. Following incubation for 24 h, cells were inactivated and fixed with 4.5% formaldehyde. Infected cells were detected using an antibody against SARS-CoV-2 NP (ThermoFischer, PA5-81794). Foci were counted using an AID ELISpot reader from Mabtech.

Code availability

Code used in this analysis has been previously published[51]. The code for forcing adherence to the Wilson distribution is included in the same repository under a GPLv3 license.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.12.378422; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Supplementary description

In the following, we discuss those compounds that did not show significant antiviral activity but for which we could determine the binding pose based on the crystal structures.

Active site, covalent

Leupeptin is a well-known cysteine protease inhibitor and was therefore included in our screening effort as a positive control[61]. Structurally, it is highly similar to calpeptin. Indeed this peptidomimetic inhibitor also forms a thiohemiacetal and occupies the substrate pocket, much like calpeptin (Figure S2A and 3E). The binding mode is identical to the recently released room-temperature structure of M^{pro} with leupeptin (PDB-ID 6XCH).

Maleate was observed covalently bound in seven structures during hit finding. In all cases maleate served as the counter ion of the applied compound. In these crystal structures the maleate, rather than the applied compound, forms a thioether with the thiol of Cys145, modifying it to succinyl-cysteine. The thiol of Cys145 undergoes a Michael-type nucleophilic attack on the C2 of maleate. A similar adduct has been described for maleate isomerase[25] as an intermediate structure in the isomerization reaction. The covalent adduct is further stabilized by hydrogen bonds to the backbone amide of Gly143 and Cys145 to the carboxylate group (C1) of succinate. The terminal carboxylate (C4) is positioned by hydrogen bonds to the side chain of Asn142 and a water-bridged hydrogen bond to the side chain of His163 (Figure S2B).

TH-302 (Evofosfamide) is covalently linked to Cys145 through nucleophilic substitution of the bromine, leading to thioether formation (Figure S2C). The other bromine-alkane chain occupies the S1 pocket while the nitro-imidazole stretches into pocket S2. The substitution of chlorine or hydroxyl for bromines in TH-302 has been demonstrated in cell culture[62]. Our mass spectrometry analysis suggested the loss of a bromine atom (Figure S3C).

Zinc pyrithione was already demonstrated to have inhibitory activity against SARS-CoV-1 M^{pro} [(63)].The pyrithione chelates the Zn^{2+} ion which coordinates the thiolate and imidazole of the catalytic dyad residues Cys145 and His41 (Figure S2D). The remaining part of the ionophore protrudes out of the active site. This tetrahedral binding mode of zinc has previously been described for other zinc-coordinating compounds in complex with HCoV-229E M^{pro} [(64)]. Interestingly, antiviral effects against a range of corona- and non-coronaviruses have already been ascribed to zinc pyrithione, although its effect had been attributed to inhibition of RNA-dependent polymerase[65]. Zinc pyrithione exhibits both antifungal and antimicrobial properties and is known in treatment of seborrheic dermatitis.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.12.378422; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Active site, non-covalent

Adrafinil mainly binds mainly through van der Waals interactions to M^{pro}. In particular, its two phenyl rings are inserted into pockets S1' and S2 (Figure S2E). A hydrogen bond is formed between the backbone amide of Cys145 and the hydroxylamine group. The side chain of Met49 is wedged between the two phenyl rings.

Fusidic acid interacts with M^{pro} mainly through hydrophobic interactions, especially through the alkene chain within pocket S2 and the tetracyclic moiety packing against Ser46 (Figure S2F). Moreover, the carboxylate group forms indirect hydrogen bonds, mediated via two water molecules, to the main chain of Thr26, Gly143 and Cys145. In addition, the same carboxylate group forms a hydrogen bond to an imidazole molecule from the crystallization conditions. This imidazole occupies pocket S1' and mediates hydrogen bonds to the backbone of His41 and Cys44. These indirect interactions offer opportunities for optimization of compounds binding to M^{pro}. Fusidic acid is a well-known bacteriostatic compound, with a steroid core structure.

LSN-2463359 binds mainly to M^{pro} by interaction of the pyridine ring with the S1 pocket (Figure S2G). Besides van der Waals interactions with the β-turn Phe140-Ser144, contributing to the pocket, it also forms a hydrogen bond to the side chain of His163.

SEN1269 binds only to the active site of one protomer in the native dimer. This causes a break in the crystallographic symmetry, leading to a different crystallographic space group (Suppl Table S2). The central pyrazine ring forms a hydrogen bond to Gln189 (Figure S2H). The terminal dimethylaniline moiety sits deep in pocket S2 which is enlarged by an outwards movement of the short α-helix Ser46-Leu50 by 1.7 Å (Ser46 Cα-atom) compared to the native structure. This includes a complete reorientation of the side chain of Met49 which now points outside of the S2 pocket. Additionally, the C-terminus of a crystallographic neighboring M^{pro} protomer is trapped between SEN1269 and part of the S1 pocket, including a hydrogen bond between Asn142 and the backbone amide of Phe305 and Gln306 of the C-terminus.

Tretazicar binds at the active site entrance at pocket S3/S4 (Figure S2I). The amide group forms hydrogen bonds to the backbone carbonyl of Glu166, the adjacent nitro group forms hydrogen bonds to the side chain of Gln192 and the backbone amide of Thr190.

UNC2327 binds to active site of M^{pro} by stacking its benzothiadiazole ring against the loop Glu166-Pro168 that forms the shallow pocket S3 (Figure S2J). This is stabilized by a hydrogen bond between the benzothiadiazole and the main chain carbonyl of Glu166. The piperine ring and adjacent carbonyl are inserted into pocket S1' and interact with Thr25 and His41.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.12.378422; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Covalent binder to Cys156

Aurothioglucose

In the crystal structure of the aurothioglucose complex, the strong nucleophile Cys145 becomes oxidized to a sulfinic acid. The initial reaction is the disproportionation of Aurothioglucose into Au(0) and a disulfide dimer of thioglucose. This is followed by a cascade of redox reactions of thioglucose, its disulfide and sulfenic acid. A disulfide linkage to thioglucose is only observed at Cys156 on the surface of M^{pro} (Figure S2K). Here the thioglucose moiety is located between Lys100 and Lys102.

Glutathione isopropyl ester binds to the surface-exposed Cys156 via a disulfide linkage (Figure S2L). Additionally, the ester forms a hydrogen bond to the backbone amide of Tyr101, while the amine of the other arm of the molecule is interacting with the side chain amine of Lys102.

Distal pockets

AR-42 binds with its phenyl ring to a small hydrophobic pocket in the dimerization domain formed by residues Gly275, Met276, Leu286 and Leu287 (Figure S2M). Additionally, the central amide forms a hydrogen bond to the backbone carbonyl of Leu272.

AZD6482 binds to a pocket on the back of the catalytic domain, away from the native dimer interface (Figure S2N). The nitrobenzene ring is inserted in a pocket formed by His80, Lys88, Leu89 and Lys90. The central aromatic system and morpholine ring lie flat on the surface of M^{pro}. Furthermore, Asn63 forms a hydrogen bond to the keto-group in the pyrimidine ring.

Climbazole binds in a shallow surface pocket, wedged between two crystallographic symmetry-related molecules (Figure S2O). Only van der Waals interactions are observed. One monomer contributes with residues Phe103, Val104, Arg105 and Glu178 to this binding site, while the other monomer contributes Asn228, Asn231, Leu232, Met235 and Pro241.

Clonidine also sits in between two crystallographic, symmetry-related molecules and binds through van der Waals interactions (Figure S2P). Here one protomer mainly forms the binding site, by contributing Asp33, Aps34 and Ala94. The other protomer contributes Lys236, Tyr237 and Asn238. The amine ring of clonidine forms a loose ring stacking interaction to Tyr237, while a hydrogen bond between the backbone carbonyl of Lys236 and the ring connecting amine of clonidine is formed. The side chain of Lys236 is flipped to the side to make room for the chlorine containing ring system.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.12.378422; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Ipidacrine is in contact with two different M^{pro} protomers (Figure S2Q). The tricylic ring system is packed against a surface loop, including residues Pro96 and Lys97 as well as Lys12. It also interacts with the end of an α-helix including residues Gln273, Asn274 and Gly275.

Tegafur binds to a in a shallow surface pocket generated by residues Asp33, Pro99, Lys100 and Tyr101. The main interaction is through π-stacking of the aromatic ring of Tyr223. The side chain of Lys100 flips away and generates space for the compound (Figure S2R).

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.12.378422; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

References

1. Ellinger, B. *et al.* 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.
2. Liu, Y. *et al.* The development of Coronavirus 3C-Like protease (3CLpro) inhibitors from 2010 to 2020. *European Journal of Medicinal Chemistry* **206**, 112711 (2020).
3. Pushpakom, S. *et al.* Drug repurposing: progress, challenges and recommendations. *Nature Reviews Drug Discovery* **18**, 41–58 (2019).
4. Zhang, L. *et al.* Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved α-ketoamide inhibitors. *Science* **368**, 409–412 (2020).
5. Jin, Z. *et al.* Structure of M pro from COVID-19 virus and discovery of its inhibitors. *Nature* 1–9 (2020) doi:10.1038/s41586-020-2223-y.
6. Hilgenfeld, R. From SARS to MERS: crystallographic studies on coronaviral proteases enable antiviral drug design. *The FEBS Journal* **281**, 4085–4096 (2014).
7. Anderson, A. C. The Process of Structure-Based Drug Design. *Chemistry & Biology* **10**, 787–797 (2003).
8. Batool, M., Ahmad, B. & Choi, S. A Structure-Based Drug Discovery Paradigm. *Int J Mol Sci* **20**, (2019).
9. Erlanson, D. A., Fesik, S. W., Hubbard, R. E., Jahnke, W. & Jhoti, H. Twenty years on: the impact of fragments on drug discovery. *Nature Reviews Drug Discovery* **15**, 605–619 (2016).
10. Burkhardt, A. *et al.* Status of the crystallography beamlines at PETRA III. *Eur. Phys. J. Plus* **131**, 56 (2016).
11. Grimes, J. M. *et al.* Where is crystallography going? *Acta Cryst D* **74**, 152–166 (2018).
12. Corsello, S. M. *et al.* The Drug Repurposing Hub: a next-generation drug library and information resource. *Nat Med* **23**, 405–408 (2017).
13. Hann, M. M., Leach, A. R. & Harper, G. Molecular Complexity and Its Impact on the Probability of Finding Leads for Drug Discovery. *J. Chem. Inf. Comput. Sci.* **41**, 856–864 (2001).
14. Ehrmann, F. R. *et al.* Soaking suggests “alternative facts”: Only co-crystallization discloses major ligand-induced interface rearrangements of a homodimeric tRNA-binding protein indicating a novel mode-of-inhibition. *PLOS ONE* **12**, e0175723 (2017).
15. Ginn, H. M. Pre-clustering data sets using cluster4x improves the signal-to-noise ratio of high-throughput crystallography drug-screening analysis. *Acta Cryst D* **76**, (2020).
16. Pearce, N. M. *et al.* A multi-crystal method for extracting obscured crystallographic states from conventionally uninterpretable electron density. *Nature Communications* **8**, 15123 (2017).
17. Xue, X. *et al.* Structures of Two Coronavirus Main Proteases: Implications for Substrate Binding and Antiviral Drug Design. *Journal of Virology* **82**, 2515–2527 (2008).

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.12.378422; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

18. Altmeyer, M. *et al.* Beta-aminoketones as prodrugs for selective irreversible inhibitors of type-1 methionine aminopeptidases. *Bioorganic & Medicinal Chemistry Letters* **24**, 5310–5314 (2014).
19. Drebes, J., Kunz, M., Pereira, C. A. & Wrenger, C. B. and C. MRSA Infections: From Classical Treatment to Suicide Drugs. *Current Medicinal Chemistry* http://www.eurekaselect.com/118131/article (2014).
20. Quasthoff, S., Möckel, C., Zieglgänsberger, W. & Schreibmayer, W. Tolperisone: A Typical Representative of a Class of Centrally Acting Muscle Relaxants with Less Sedative Side Effects. *CNS Neurosci Ther* **14**, 107–119 (2008).
21. Mogi, M., Ito, T., Matsuki, Y., Kurata, Y. & Nambara, T. Determination of isofloxythepin in biological fluids by gas chromatography-mass spectrometry. *J Chromatogr* **399**, 234–250 (1987).
22. Lau, Y.-S., Fung, Y. K. & Anderson, T. M. Antagonistic effects of isofloxythepin on dopamine D1 and D2 receptors and behaviors in rats. *General Pharmacology: The Vascular System* **29**, 729–736 (1997).
23. Turk, V. *et al.* Cysteine cathepsins: From structure, function and regulation to new frontiers. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **1824**, 68–88 (2012).
24. Piccart, M. *et al.* Phase I clinical trial with alpha 1,3,5-triglycidyl-s-triazinetrione (NSC-296934). *European Journal of Cancer and Clinical Oncology* **17**, 1263–1266 (1981).
25. Fisch, F. *et al.* A Covalent Succinylcysteine-like Intermediate in the Enzyme-Catalyzed Transformation of Maleate to Fumarate by Maleate Isomerase. *Journal of the American Chemical Society* **132**, 11455–11457 (2010).
26. Ma, C. *et al.* Boceprevir, GC-376, and calpain inhibitors II, XII inhibit SARS-CoV-2 viral replication by targeting the viral main protease. *Cell Research* 1–15 (2020) doi:10.1038/s41422-020-0356-z.
27. Vuong, W. *et al.* Feline coronavirus drug inhibits the main protease of SARS-CoV-2 and blocks virus replication. *Nature Communications* **11**, 4282 (2020).
28. Sacco, M. D. *et al.* Structure and inhibition of the SARS-CoV-2 main protease reveals strategy for developing dual inhibitors against Mpro and cathepsin L. *Science Advances* eabe0751 (2020) doi:10.1126/sciadv.abe0751.
29. Escaich, S. *et al.* The MUT056399 Inhibitor of FabI Is a New Antistaphylococcal Compound. *Antimicrobial Agents and Chemotherapy* **55**, 4692–4697 (2011).
30. Wissner, A. *et al.* Synthesis and Structure–Activity Relationships of 6,7-Disubstituted 4-Anilinoquinoline-3-carbonitriles. The Design of an Orally Active, Irreversible Inhibitor of the Tyrosine Kinase Activity of the Epidermal Growth Factor Receptor (EGFR) and the Human Epidermal Growth Factor Receptor-2 (HER-2). *J. Med. Chem.* **46**, 49–63 (2003).
31. Erlichman, C. *et al.* Phase I Study of EKB-569, an Irreversible Inhibitor of the Epidermal Growth Factor Receptor, in Patients With Advanced Solid Tumors. *JCO* **24**, 2252–2260 (2006).

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.12.378422; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

32. Avenet, P. *et al.* Antagonist properties of the stereoisomers of ifenprodil at NR1A/NR2A and NR1A/NR2B subtypes of the NMDA receptor expressed in Xenopus oocytes. *European Journal of Pharmacology* **296**, 209–213 (1996).
33. Mirzadegan, T. *et al.* Identification of the Binding Site for a Novel Class of CCR2b Chemokine Receptor Antagonists BINDING TO A COMMON CHEMOKINE RECEPTOR MOTIF WITHIN THE HELICAL BUNDLE. *J. Biol. Chem.* **275**, 25562–25571 (2000).
34. Belliotti, T. R. *et al.* Isoindolinone enantiomers having affinity for the dopamine D4 receptor. *Bioorganic & Medicinal Chemistry Letters* **8**, 1499–1502 (1998).
35. Shi, J., Sivaraman, J. & Song, J. Mechanism for Controlling the Dimer-Monomer Switch and Coupling Dimerization to Catalysis of the Severe Acute Respiratory Syndrome Coronavirus 3C-Like Protease. *Journal of Virology* **82**, 4620–4629 (2008).
36. Wyatt, P. G. *et al.* Identification of N-(4-piperidinyl)-4-(2,6-dichlorobenzoylamino)-1H-pyrazole-3-carboxamide (AT7519), a novel cyclin dependent kinase inhibitor using fragment-based X-ray crystallography and structure based drug design. *J Med Chem* **51**, 4986–4999 (2008).
37. Tan, J. *et al.* pH-dependent Conformational Flexibility of the SARS-CoV Main Proteinase (Mpro) Dimer: Molecular Dynamics Simulations and Multiple X-ray Structure Analyses. *Journal of Molecular Biology* **354**, 25–40 (2005).
38. Zhang, C. *et al.* Ifenprodil and Flavopiridol Identified by Genomewide RNA Interference Screening as Effective Drugs To Ameliorate Murine Acute Lung Injury after Influenza A H5N1 Virus Infection. *mSystems* **4**, (2019).
39. Douangamath, A. *et al.* Crystallographic and electrophilic fragment screening of the SARS-CoV-2 main protease. *Nature Communications* **11**, 5047 (2020).
40. Riva, L. *et al.* Discovery of SARS-CoV-2 antiviral drugs through large-scale compound repurposing. *Nature* **586**, 113–119 (2020).
41. Chen, C. Z. *et al.* Drug Repurposing Screen for Compounds Inhibiting the Cytopathic Effect of SARS-CoV-2. *bioRxiv* 2020.08.18.255877 (2020) doi:10.1101/2020.08.18.255877.
42. Drayman, N. *et al.* Drug repurposing screen identifies masitinib as a 3CLpro inhibitor that blocks replication of SARS-CoV-2 in vitro. *bioRxiv* 2020.08.31.274639 (2020) doi:10.1101/2020.08.31.274639.
43. D'Arcy, A., Bergfors, T., Cowan-Jacob, S. W. & Marsh, M. Microseed matrix screening for optimization in protein crystallization: what have we learned? *Acta Cryst F* **70**, 1117–1126 (2014).
44. Kabsch, W. XDS. *Acta Cryst D* **66**, 125–132 (2010).
45. Vonrhein, C. *et al.* Data processing and analysis with the autoPROC toolbox. *Acta Cryst D* **67**, 293–302 (2011).
46. Tickle, I. J. *et al.* STARANISO. Cambridge, United Kingdom: Global Phasing Ltd. (2018).
47. Beilsten-Edmands, J. *et al.* Scaling diffraction data in the DIALS software package: algorithms and new approaches for multi-crystal scaling. *Acta Cryst D* **76**, 385–399 (2020).

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.12.378422; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

48. Winter, G. xia2: an expert system for macromolecular crystallography data reduction. *J Appl Cryst* **43**, 186–190 (2010).
49. Liebschner, D. *et al.* Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Cryst D* **75**, 861–877 (2019).
50. https://stash.desy.de/login?next=/projects/X2P/repos/xia2pipe/browse.
51. Ginn, H. M. Pre-clustering data sets using cluster4x improves the signal-to-noise ratio of high-throughput crystallography drug-screening analysis. *Acta Cryst D* **76**, 1134–1144 (2020).
52. Murshudov, G. N. *et al.* REFMAC5 for the refinement of macromolecular crystal structures. *Acta Cryst D* **67**, 355–367 (2011).
53. Turk, D. MAIN software for density averaging, model building, structure refinement and validation. *Acta Cryst D* **69**, 1342–1357 (2013).
54. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486–501 (2010).
55. Sommer, K. *et al.* UNICON: A Powerful and Easy-to-Use Compound Library Converter. *J. Chem. Inf. Model.* **56**, 1105–1111 (2016).
56. Bietz, S. & Rarey, M. SIENA: Efficient Compilation of Selective Protein Binding Site Ensembles. *J. Chem. Inf. Model.* **56**, 248–259 (2016).
57. Bietz, S., Urbaczek, S., Schulz, B. & Rarey, M. Protoss: a holistic approach to predict tautomers and protonation states in protein-ligand complexes. *Journal of Cheminformatics* **6**, 12 (2014).
58. Flachsenberg, F., Meyder, A., Sommer, K., Penner, M. & Rarey, M. A Consistent Scheme for Gradient-Based Optimization of Protein-Ligand Poses. *ACS Journal of Chemical Information and Modeling* (2020).
59. Schneider, N., Lange, G., Hindle, S., Klein, R. & Rarey, M. 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).
60. van den Heuvel, R. H. H. *et al.* Improving the Performance of a Quadrupole Time-of-Flight Instrument for Macromolecular Mass Spectrometry. *Anal. Chem.* **78**, 7473–7483 (2006).
61. Aoyagi, T. *et al.* BIOLOGICAL ACTIVITIES OF LEUPEPTINS. *J. Antibiot.* **22**, 558–568 (1969).
62. Hong, C. R. *et al.* Cellular pharmacology of evofosfamide (TH-302): A critical re-evaluation of its bystander effects. *Biochemical Pharmacology* **156**, 265–280 (2018).
63. Hsu, J. T.-A. *et al.* Evaluation of metal-conjugated compounds as inhibitors of 3CL protease of SARS-CoV. *FEBS Letters* **574**, 116–120 (2004).
64. Lee, C.-C. *et al.* Structural Basis of Inhibition Specificities of 3C and 3C-like Proteases by Zinc-coordinating and Peptidomimetic Compounds. *J. Biol. Chem.* **284**, 7646–7655 (2009).
65. Velthuis, A. J. W. te *et al.* Zn2+ Inhibits Coronavirus and Arterivirus RNA Polymerase Activity In Vitro and Zinc Ionophores Block the Replication of These Viruses in Cell Culture. *PLOS Pathogens* **6**, e1001176 (2010).

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.12.378422; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

66. Wollenhaupt, J. *et al.* F2X-Universal and F2X-Entry: Structurally Diverse Compound Libraries for Crystallographic Fragment Screening. *Structure* **28**, 694-706.e5 (2020).

67. Cox, O. B. *et al.* A poised fragment library enables rapid synthetic expansion yielding the first reported inhibitors of PHIP(2), an atypical bromodomain. *Chem. Sci.* **7**, 2322–2330 (2016).

Acknowledgements

We acknowledge DESY (Hamburg, Germany), a member of the Helmholtz Association HGF, for the provision of experimental facilities. Parts of this research were carried out at PETRA III at beamline P11. Further MX data was collected at beamline P13 and P14 operated by EMBL. This research was supported in part through the Maxwell computational resources operated at Deutsches Elektronen-Synchrotron (DESY), Hamburg, Germany. We acknowledge the use of the XBI biological sample preparation laboratory, enabled by the XBI User Consortium. We acknowledge financial support through the EXSCALATE4CoV EU-H2020 Emergency Project (https://www.exscalate4cov.eu). We also acknowledge financial support by the Cluster of Excellence 'Advanced Imaging of Matter' of the Deutsche Forschungsgemeinschaft (DFG) - EXC 2056 - project ID 390715994, the Federal Ministry of Education and Research (BMBF) via project 05K19GU4, 05K16GUA and 05K20FL1 ("STOP CORONA"), and the Joachim-Herz-Stiftung Hamburg via the project Infecto-Physics. CE and MR acknowledge the financial support by Grant-No. HIDSS-0002 DASHH (Data Science in Hamburg - HELMHOLTZ Graduate School for the Structure of Matter). The work of JPZ and MR was supported by the BMBF (SFX2: Hochdurchsatz Serielle-Femtosekunden Kristallographie@EU XFEL (WP2) - Compound Target Screenings of essential SARS-CoV-2 Enzymes and selected Human Corona processing Enzymes FKZ: 05K19GU4). DT is supported by Slovenian Research Agency (ARRS; research program P1-0048, Infrastructural program IO-0048 – both awarded to D.T.). B.S. was supported by an Exploration Grant from the Boehringer Ingelheim Foundation. R.C. is supported by DFG grants INST 187/621-1 and INST 187/686-1.

We thank Thomas Dietrich, Simon Geile, Heshmat Noei, and Tim Pakendorf from DESY, and Bianca Di Fabrizio and Sebastian Kühn from BNI for assistance. Finally, we thank the DESY machine group, in particular Mario Wunderlich, Kim Heuck, Arne Brinkmann, Olaf Goldbeck, Jürgen Haar, Torsten Schulz, Gunnar Priebe, Maximilian Holz, Björn Lemcke, Klaus Knaack, Oliver Seebauer, Philipp Willanzheimer, Rolf Jonas, and Nicole Engling.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.12.378422; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

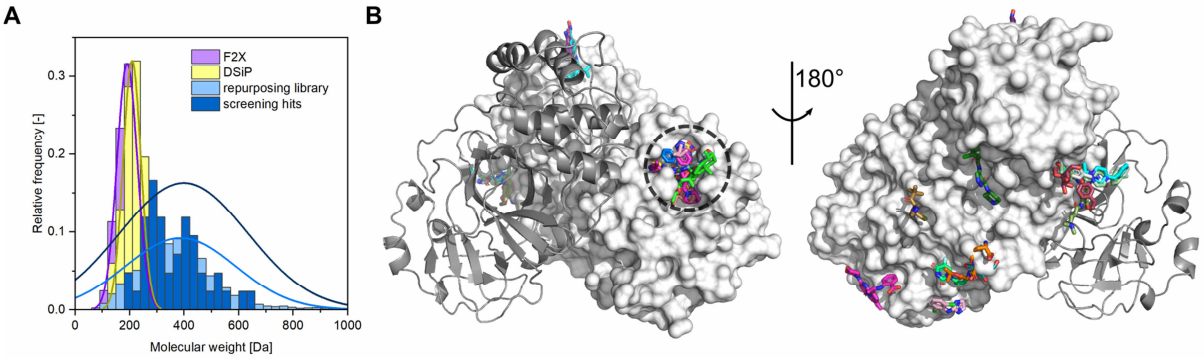


Figure 1: The repurposing libraries reveal compound binding sites distributed across the complete M^{pro} surface. **A**, Normalized histograms of molecular weight distributions of two commonly used fragment screening libraries F2X-Universal[66] (median 193.2 Da) and DSiP (a version of the "poised library"[67], median 211.2 Da), the two combined repurposing libraries used in the present effort, and the resulting hits from our X-ray screen (Fraunhofer IMG median 371.3 Da, Dompé "Safe-in-man" 316.3 Da, combined 366.5 Da). Normal distributions are indicated by solid lines in corresponding colors. Compounds with a molecular weight above 1000 Da are not shown. **B**, Cartoon representation of M^{pro} with all unambiguously bound compounds. One protomer of the native dimer is depicted as a cartoon and the other one as surface representation. Left panel, view of the active site of M^{pro}, right panel, view of M^{pro} rotated by 180°. The active site is indicated by a dashed circle.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.12.378422; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

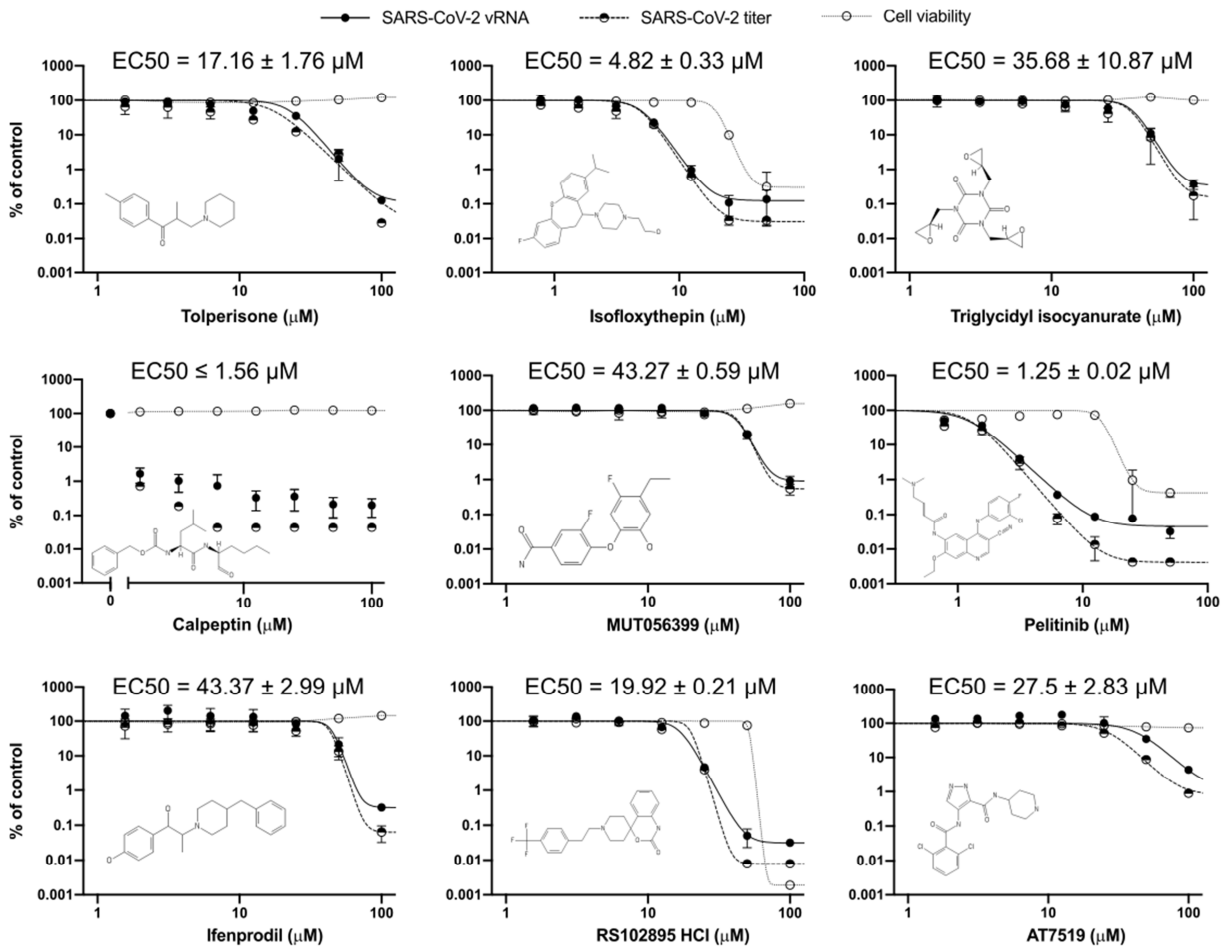


Figure 2: Effect of selected compounds on SARS-CoV-2 replication in Vero E6 cells. The vRNA yield (solid circles), viral titers (half-solid circles), and cell viability (empty circles) were determined by RT-qPCR, immunofocus assays, and the CCK-8 method, respectively. EC_{50} for the viral titers reduction are shown. Values were calculated from three independent data sets.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.12.378422; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

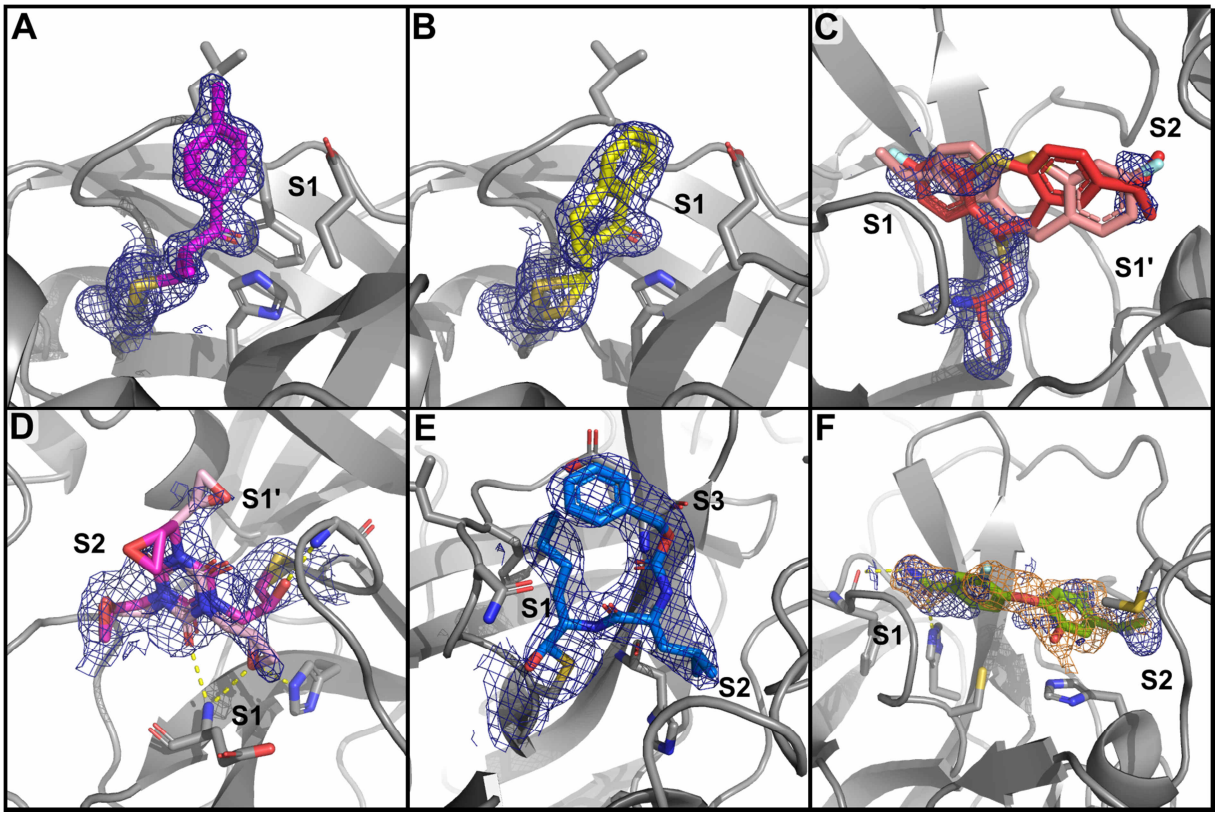


Figure 3: Detailed view of covalent and non-covalent binders in the active site of M^{pro}. Bound compounds are depicted as colored sticks while M^{pro} is shown as a grey cartoon representation with selected interacting residues as sticks. Hydrogen bonds are depicted by dashed lines. The blue mesh represents 2Fo-Fc electron-density maps carved at 1.6 Å around the compounds (rmsd = 1 except for **E** and **F**, which are shown at rmsd = 0.7). For **E** the PanDDA event map is additionally shown in orange (rmsd = 1). **A**, tolperisone; **B**, HEAT; **C**, isofloxythepin; **D**, triglycidyl isocyanurate; **E**, calpeptin; **F**, MUT056399. Additional information is provided in Suppl. Table S1.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.12.378422; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

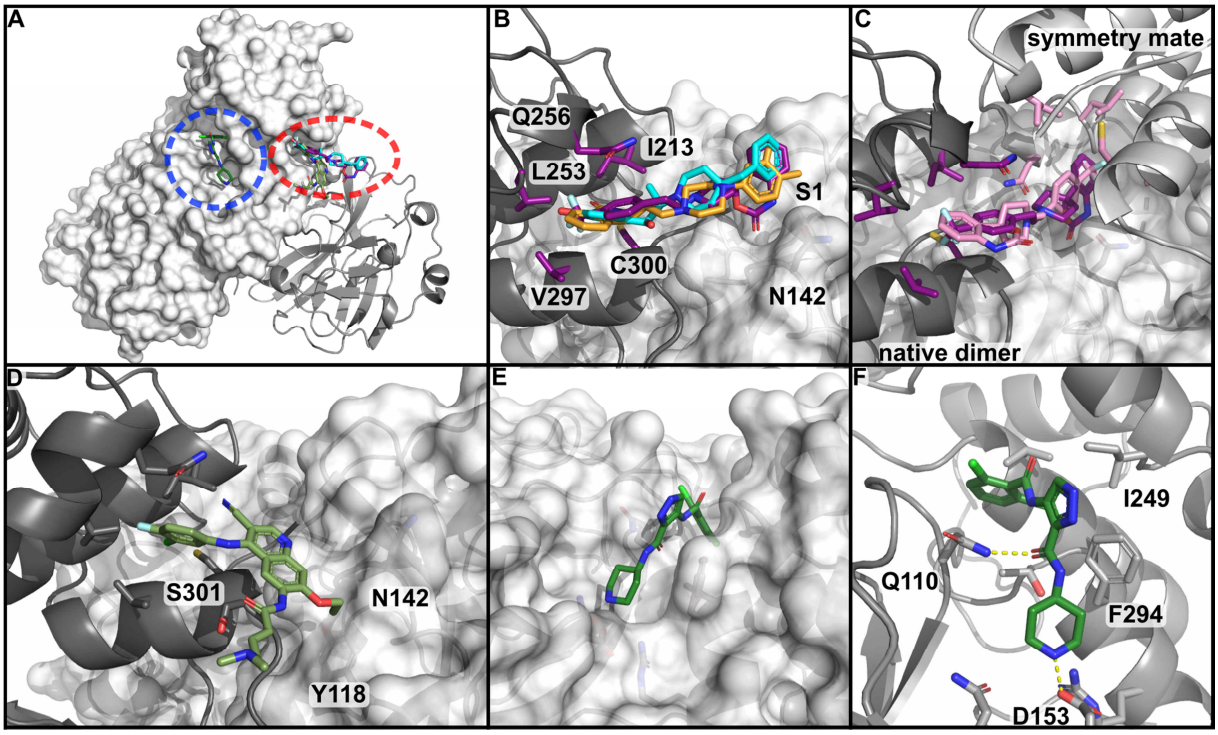
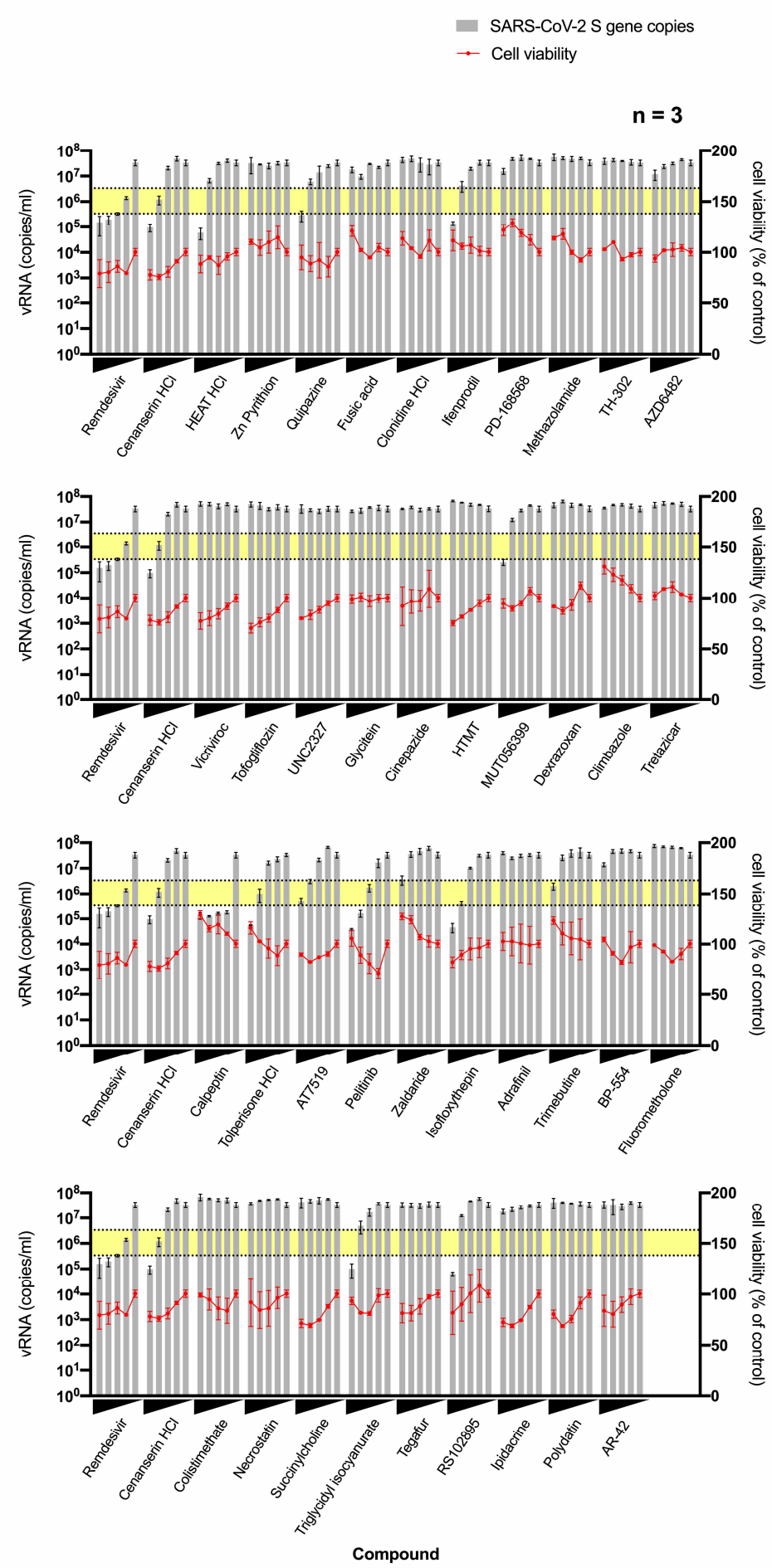


Figure 4: Screening hits at allosteric sites of M^{pro}. **A**, View of the allosteric sites of M^{pro}. One site is within the dimerization domain of M^{pro} proximal to the active site (red circle). The other site is in between the catalytic domains and the dimerization domain in a deep groove (blue). **B**, Close up view of the binding site in the dimerization domain, close to the active site of second protomer in the native dimer. Residues forming the hydrophobic pocket are indicated. RS-102895 (purple), ifenprodil (cyan) and PD-168568 (orange) cross the native dimer interface and reach the rim of pocket S1 of the active site of the other protomer. **C**, crystallographic dimer (not representing the native dimer) generates an inverted binding mode of compounds at the same binding site. **D**, Pelitinib binds to the C-terminal α-helix at Ser301 and pushes against Asn142 and the β-turn of the pocket S1. **E**, AT7519 occupies a deep cleft between the catalytic and dimerization domain of M^{pro}. **F**, M^{pro} residues interacting with the compound AT7519 are depicted as sticks, hydrogen bonds are indicated by dashed lines. Additional information is provided in Suppl. Table S1.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.12.378422; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

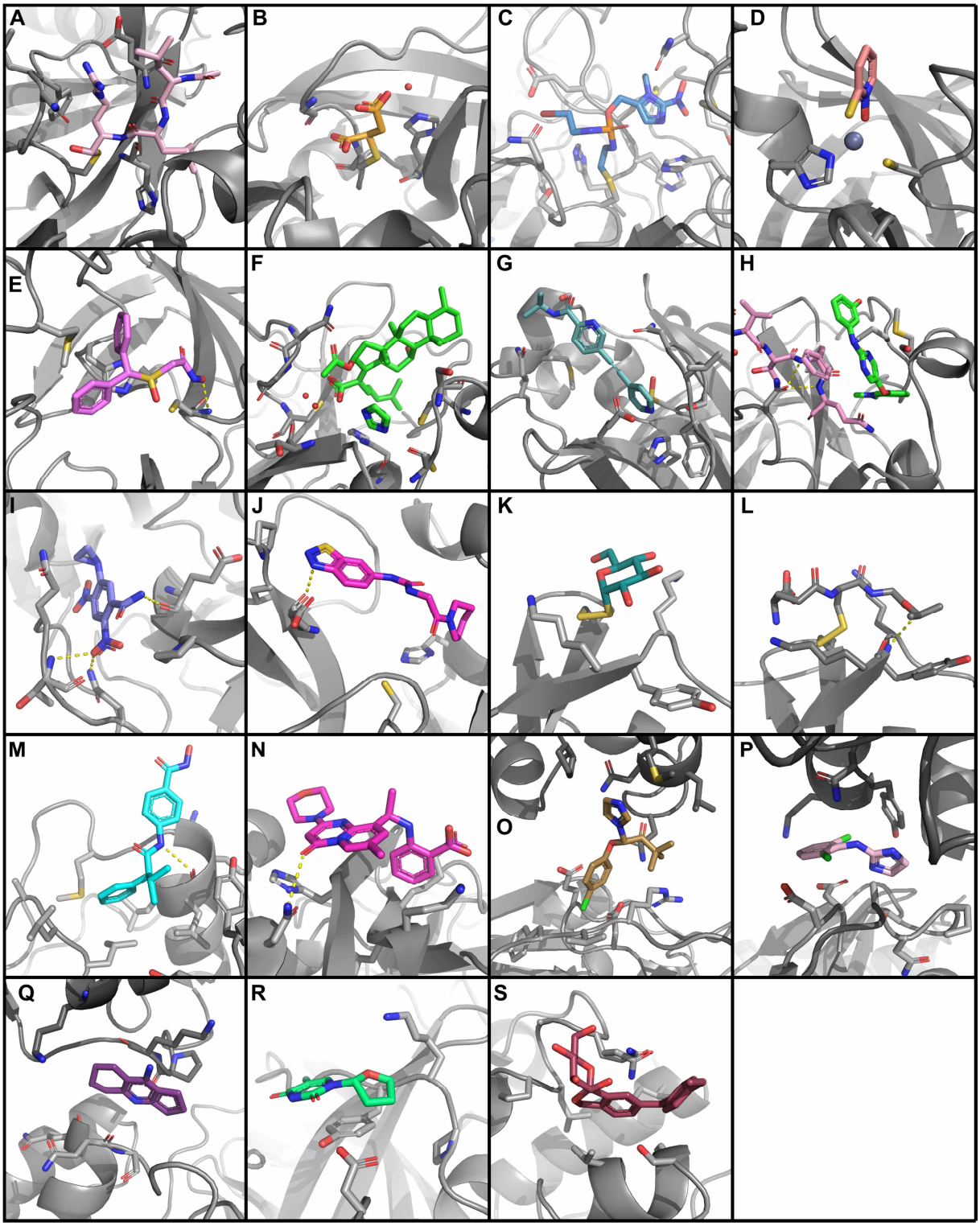


1020

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.12.378422; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Supplementary Figure S1: X-ray hit compounds were tested in a non-toxic range for inhibition of SARS-CoV-2 replication in Vero E6 cells. The vRNA yield (gray bars) and cell viability (red circles) were determined by RT-qPCR and the CCK-8 method, respectively. All data are mean ± standard deviation. Upper and lower boundaries of yellow bars represent one and two log reduction in vRNA level. Twofold serial dilutions of compounds were used to treat cells for 42 hours, where 100 µM was used as the highest concentration for all compounds except Remdesivir (10 µM), Cenanserin HCl (125 µM, HEAT HCl (25 µM), Zn Pyrithion (1 µM), Pelitinib (12.5 µM), Zaldaride (50 µM), Isofloxythepin (25 µM) and RS102895 HCl (50 µM). Control is DMSO without compound.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.12.378422; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

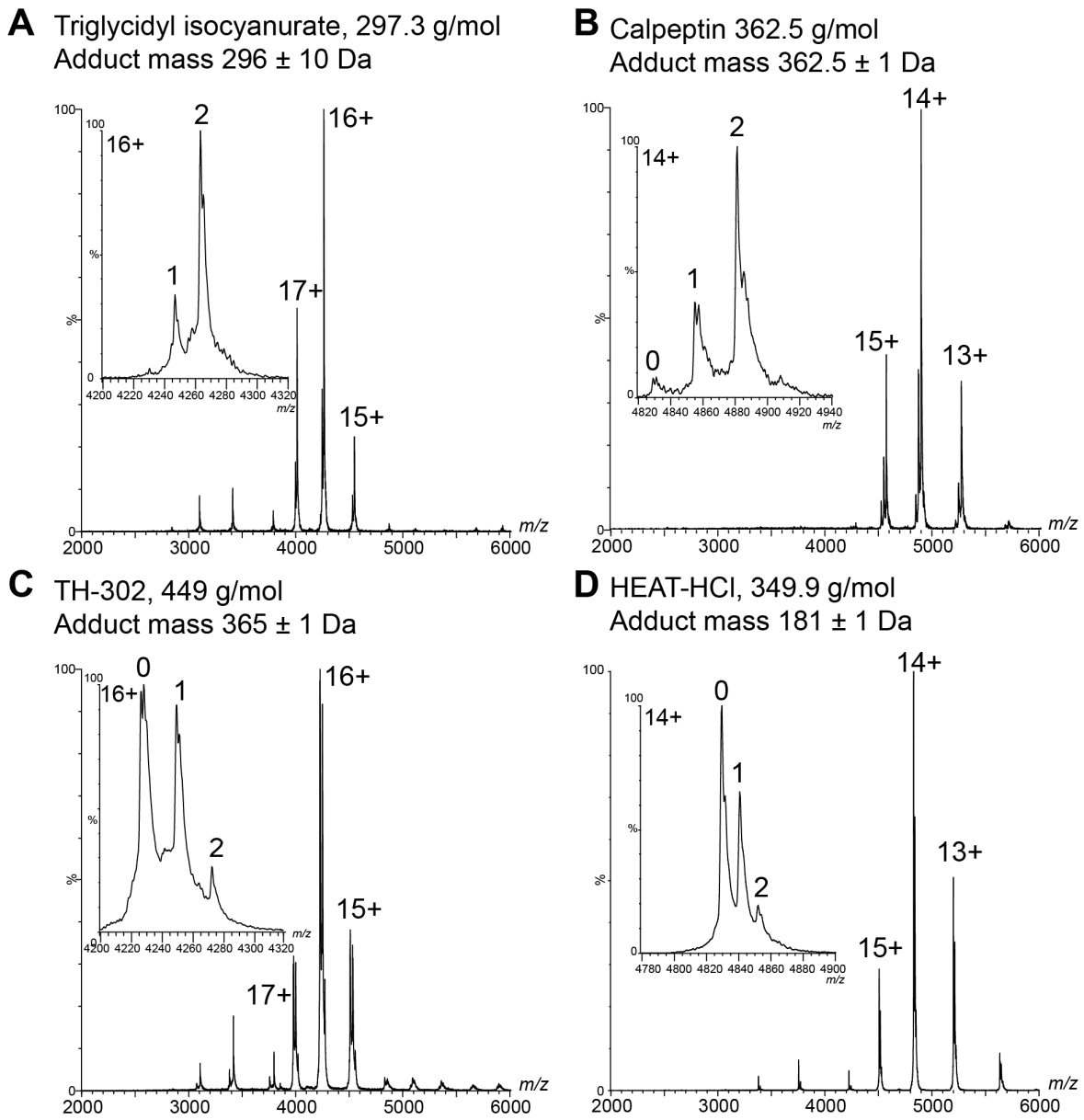


Supplementary Figure S2: The structures of inactive compounds. Compounds are depicted as colored sticks. M^{pro} is shown as a grey cartoon model with residues important for ligand binding shown as stick models and hydrogen bonds are indicated by dashed lines. Ligands binding covalently to the active site residue Cys145: **A**, leupeptin. **B**, maleate. **C**, TH-302. **D**, zinc pyrithione. Ligands binding non-covalently to the active site: **E**, adrafinil. **F**, fusidic acid. **G**, LSN-2463359. **H**, SEN1269 (C-terminus of neighboring M^{pro} protomer shown as pink stick model). **I**, tretazicar. **J**, UNC2327. Covalent binders to Cys156: **K**, aurothioglucose. **L**, glutathione

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.12.378422; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

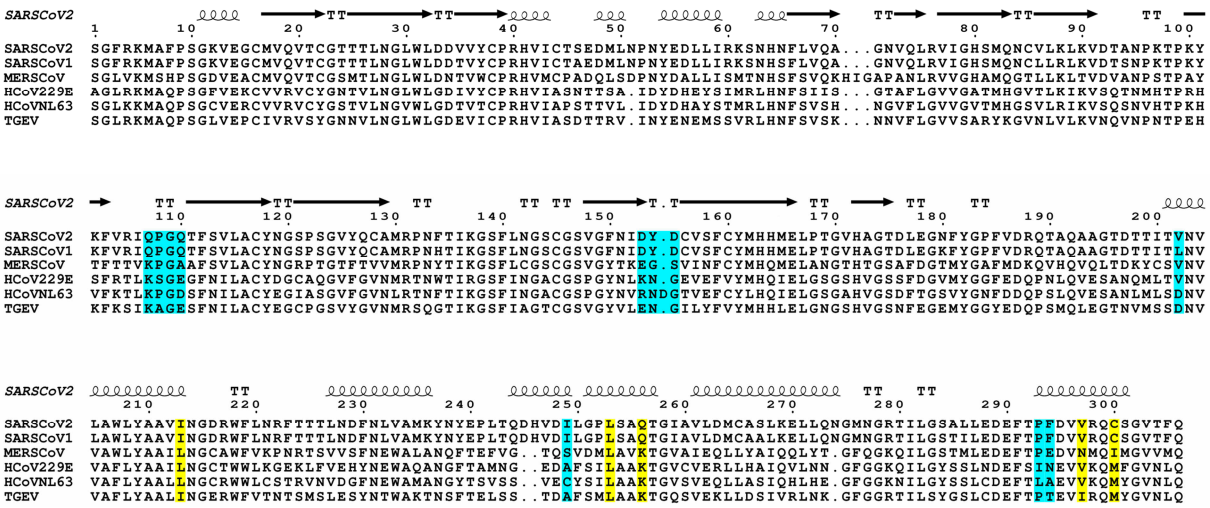
isopropylester. Other surface pockets: **M**, AR-42. **N**, AZD6482. **O**, climbazole. **P**, clonidine. **Q**, ipidacrine. **R**, tegafur. Allosteric binding site: **S**, tofogliflozin.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.12.378422; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



Supplementary Figure S3: Binding of compounds confirmed by native mass-spectrometry. Main mass spectra of M^{pro} with compounds. **(A)**, Triglycidyl isocyanurate, **(B)** calpeptin, **(C)** TH-302 and **(D)** HEAT-HCl. Insets depict main charge state signals with native M^{pro} (0) binding to one (1) or two (2) compounds, exhibiting the molecular mass of the complete compound (A and B) or a fragment (C and D). Mass spectra were recorded after the inhibitor was washed out (A and C) or in presence of fivefold excess of compound (B and D). Average compound masses are given and charged states are labelled.

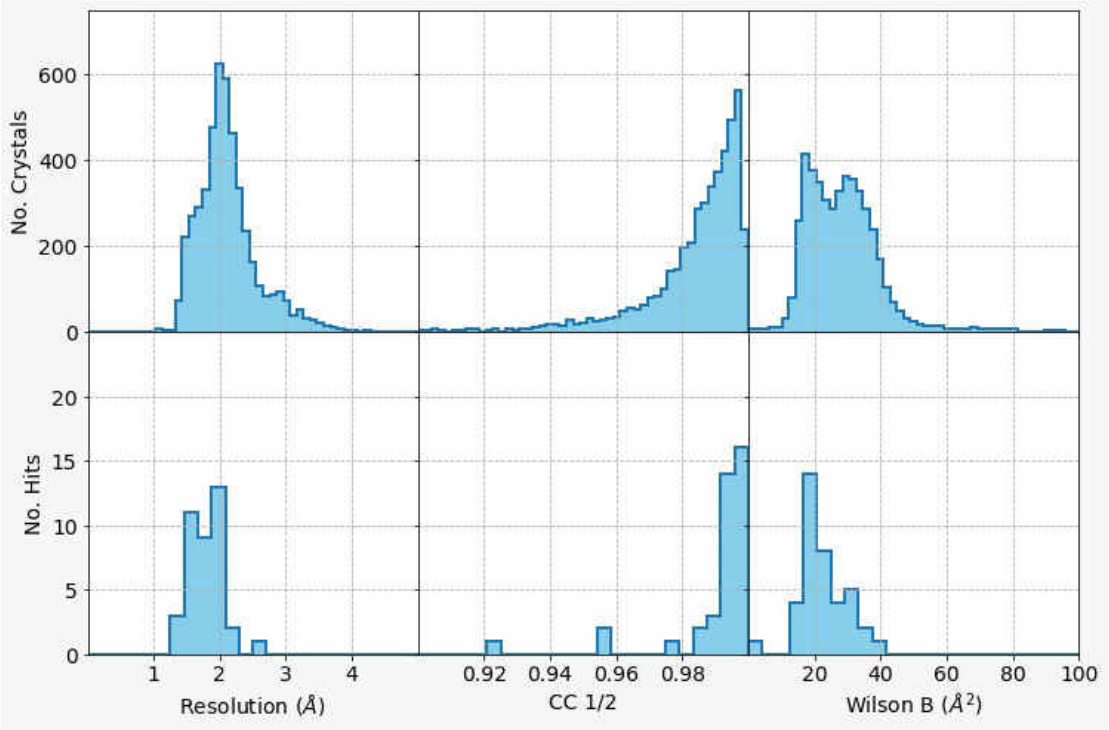
bioRxiv preprint doi: https://doi.org/10.1101/2020.11.12.378422; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



Supplementary Figure S4: Sequence comparison of M^{pro} of α- (HCoV229E, HCovNL63, TGEV) and β- (SARS-CoV-1&2, MERS-CoV) coronaviruses. SARS-CoV-2 M^{pro}-residues interacting with compounds in allosteric site 1 (blue) and 2 (yellow) are indicated in colored boxes.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.12.378422; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Supplementary Figures



Supplementary Figure S5: Distribution of data quality indicators of all collected X-ray diffraction datasets (upper panel) and of datasets with identified compound (lower panel): diffraction resolution (left), CC1/2 of the datasets (middle), and Wilson B-factor (right).

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.12.378422; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Supplementary Table S3: Native MS verified binding of compounds to M^{pro}. The table shows compounds and their molecular weight. Mass spectra of compounds and M^{pro} (final conc. 50 µM and 10 µM) were analyzed by converting peak intensities into intensity fractions for zero, one and two ligands (0/1/2 ligands in %) bound per M^{pro} dimer. Mass of the fragmented compounds is given when the observed mass is deviating from the expected mass.

Compound	Compound mass [Da]	Intensity fraction (0/1/2 compounds per M^{pro} dimer) [%]	Mass of fragmented compounds [Da]
Calpeptin	362.5	5 / 27 / 68	
Triglycidyl isocyanurate	297.3	8 / 38 / 53	
Zinc Pyrithione	317.7	11 / 34 / 55	128
HEAT HCl	349.9	39 / 37 / 24	181
HTMT	614.6	55 / 24 / 21	131
Dexrazoxan	268.3	56 / 29 / 15	
Adrafinil	289.4	58 / 28 / 14	
TH-302 (Evofosfamide)	449.0	59 / 30 / 10	365
Ifenprodil	325.2	61 / 24 / 15	126
AZD6482	408.5	62 / 29 / 9	
Glutathione-monoisopropyl-ester	349.4	62 / 28 / 10	126/188
AT7519	382.2	65 / 28 / 7	
AL-8697	402.4	67 / 24 / 8	
Cinepazide maleate	533.6	72 / 21 / 7	
UNC2327	319.4	76 / 24 / 0	
Fusidic acid	516.7	78 / 18 / 5	
AR-42	312.4	78 / 18 / 4	580
PD-168568 (HCl)2	440.4	81 / 16 / 3	350
Tofogliflozin (hydrate)	404.5	82 / 18 / 0	380
MUT056399	293.3	83 / 17 / 0	
Colistimethate Na	1735.8	83 / 15 / 2	
Vicriviroc (maleate)	649.7	88 / 12 / 0	535
Pelitinib	467.9	88 / 12 / 0	

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.12.378422; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Supplementary Table Legends

Supplementary Table S1: Comprehensive summary sheets of hit compounds showing electron-density maps, compound interactions with M^{pro}, detailed compound information, biochemical and cell-based antiviral reduction data.

Supplementary Table S2: Summary of X-ray crystallographic data processing and refinement statistics.

Supplementary Table S4: The highest ranked 200 compounds of the virtual screening. The names and HYDE scores of the top ranked molecules are given. The yellow background highlights compounds for which high-quality X-ray data was obtained in the X-ray screening. The green background highlights compounds that were detected in the active site in the X-ray screen. Compounds highlighted in light green show a similar binding mode to the fragment with the PDB ligand ID K0G in complex with M^{pro} (PDB ID 5R83). Compounds highlighted in light yellow were reported as being active in other screening studies.