X-ray screening identifies active site and allosteric inhibitors of SARS-CoV-2 main protease

Sebastian Günther1,1, Patrick Y. A. Reinke1, Yaiza Fernández-García2, Julia Lieske3, Thomas J. Lane3, Helen M. Ginn3, Faisal H. M. Koua1, Christiane Ehrt4, Wiebke Ewert1, Dominik Oberthuer1, Oleksandr Yefanov1, Susanne Meier5,6, Kristina Lorenzen7, Boris Kirchel1, Janine-Denise Kopicki8, Luca Gelisio1, Wolfgang Brehm1, Ilona Dunkel9, Brenna Norton-Baker11,12, Beatriz Escudero-Pérez2, Martin Domaracky1, Sofiane Saouane13, Amanda Tolstikova1,11, Henry Gieseler5,6, Hina Andaleeb16, Alexandra Tolstikova11, Brandon Seychell10, Thomas A. White1, Anna Hänle1, Michael Groessler1, Huijong Han7, Holger Fleckenstein1, Stephan Niebling19, Christian Günther19, Christina Schmidt7, Robin Schubert7, Gisel E. Peña-Murillo1, Aida Rahmani Mashhour1, Vincent Hennicke1, Pontus Fischer1, Luca Gelisio1, Wolfgang Brehm1, Ilona Dunkel9, Brenna Norton-Baker11,12, Beatriz Escudero-Pérez2, Martin Domaracky1, Sofiane Saouane13, Amanda Tolstikova11, Brandon Seychell10, Thomas A. White1, Anna Hänle1, Michael Groessler1, Huijong Han7, Holger Fleckenstein1, Stephan Niebling19, Christian Günther19, Christina Schmidt7, Robin Schubert7, Hjui Jang9, Juliane Boger20, Diana C. F. Monteiro21, Linlin Zhang20,22, Xinyuanyuan Sun20,22, Russell Cox28, Andrea Zaliani17, Tobias Beck5,10, Matthias Rarey4, Stephan Günther2, Dusan Turk24,25, Henning Tidow5,26, Ashwin Chari27, Rolf Hilgenfeld20,22, Charlotte Uetrecht8, Saravanan Panneerselvam19, Ivars Karpics19, Thomas R. Schneider19, Maria Marta Garcia-Alai19, Oleksandr Yefanov1, Susanne Meier5,6, Kristina Lorenzen7, Boris Krichel8, Janine-Denise Kopicki8, Helen M. Ginn3, Faisal H. M. Koua1, Christiane Ehrt4, Wiebke Ewert1, Dominik Oberthuer1, Oleksandr Yefanov1, Susanne Meier5,6, Kristina Lorenzen7, Boris Kirchel1, Janine-Denise Kopicki8, Luca Gelisio1, Wolfgang Brehm1, Ilona Dunkel9, Brenna Norton-Baker11,12, Beatriz Escudero-Pérez2, Martin Domaracky1, Sofiane Saouane13, Amanda Tolstikova11, Brandon Seychell10, Thomas A. White1, Anna Hänle1, Michael Groessler1, Huijong Han7, Holger Fleckenstein1, Stephan Niebling19, Christian Günther19, Christina Schmidt7, Robin Schubert7, Gisel E. Peña-Murillo1, Aida Rahmani Mashhour1, Vincent Hennicke1, Pontus Fischer1, Luca Gelisio1, Wolfgang Brehm1, Ilona Dunkel9, Brenna Norton-Baker11,12, Beatriz Escudero-Pérez2, Martin Domaracky1, Sofiane Saouane13, Amanda Tolstikova11, Brandon Seychell10, Thomas A. White1, Anna Hänle1, Michael Groessler1, Huijong Han7, Holger Fleckenstein1, Stephan Niebling19, Christian Günther19, Christina Schmidt7, Robin Schubert7, Hjui Jang9, Juliane Boger20, Diana C. F. Monteiro21, Linlin Zhang20,22, Xinyuanyuan Sun20,22, Russell Cox28, Andrea Zaliani17, Tobias Beck5,10, Matthias Rarey4, Stephan Günther2, Dusan Turk24,25, Henning Tidow5,26, Ashwin Chari27, Rolf Hilgenfeld20,22, Charlotte Uetrecht8, Saravanan Panneerselvam19, Ivars Karpics19, Thomas R. Schneider19, Maria Marta Garcia-Alai19, Oleksandr Yefanov1, Susanne Meier5,6, Kristina Lorenzen7, Boris Krichel8, Janine-Denise Kopicki8, Helen M. Ginn3, Faisal H. M. Koua1, Christiane Ehrt4, Wiebke Ewert1, Dominik Oberthuer1, Oleksandr Yefanov1, Susanne Meier5,6, Kristina Lorenzen7, Boris Kirchel1, Janine-Denise Kopicki8, Luca Gelisio1, Wolfgang Brehm1, Ilona Dunkel9, Brenna Norton-Baker11,12, Beatriz Escudero-Pérez2, Martin Domaracky1, Sofiane Saouane13, Amanda Tolstikova11, Brandon Seychell10, Thomas A. White1, Anna Hänle1, Michael Groessler1, Huijong Han7, Holger Fleckenstein1, Stephan Niebling19, Christian Günther19, Christina Schmidt7, Robin Schubert7, Hjui Jang9, Juliane Boger20, Diana C. F. Monteiro21, Linlin Zhang20,22, Xinyuanyuan Sun20,22, Russell Cox28, Andrea Zaliani17, Tobias Beck5,10, Matthias Rarey4, Stephan Günther2, Dusan Turk24,25, Henning Tidow5,26, Ashwin Chari27, Rolf Hilgenfeld20,22, Charlotte Uetrecht8, Saravanan Panneerselvam19, Ivars Karpics19, Thomas R. Schneider19, Maria Marta Garcia-Alai19, Stephan Niebling19, Christian Günther19, Christina Schmidt7, Robin Schubert7, Hjui Jang9, Juliane Boger20, Diana C. F. Monteiro21, Linlin Zhang20,22, Xinyuanyuan Sun20,22, Russell Cox28, Andrea Zaliani17, Tobias Beck5,10, Matthias Rarey4, Stephan Günther2, Dusan Turk24,25, Winfried Hinrichs16,29, Henry N. Chapman15,30, Arwen R. Pearson5,6, Christian Betz5,6, Alke Meents11

The coronavirus disease (COVID-19) caused by SARS-CoV-2 is creating tremendous human suffering. To date, no effective drug is available to directly treat the disease. In a search for a drug against COVID-19, we have performed a high-throughput x-ray crystallographic screen of two repurposing drug libraries against SARS-CoV-2 main protease (Mpro), which is essential for viral replication. In contrast to commonly large polyproteins which must be separated frames are expressed as two overlapping 30,000 nucleotides. The viral open reading age is primarily accomplished by the main transcription activity (1). Mpro possesses a chymotrypsin-like fold (2). Mpro is catalytically active and contains a chymotrypsin-like fold with a catalytic dyad comprised of Cys145 and His41 in its active site, which is formed by four major pockets that are labeled according to their position relative to the scissile bond of the substrate (Fig. 1) (1). The active site is located in a cleft between the two N-terminal domains of the three-domain structure of the monomer, whereas the C-terminal helical domain is involved in regulation and dimerization of the enzyme (Fig. 1A). Because of its central involvement in viral infection, Mpro 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 (3). Indeed, a number of recent publications confirm the potential of targeting Mpro for inhibition of viral replication (1, 2, 4).

In order to find drug candidates against SARS-CoV-2, we performed a large-scale x-ray crystallographic screen of Mpro against two repurposing libraries containing 5953 compounds from the Fraunhofer IME Repurposing Collection and the Safe-in-man library from Dompé Farmaceutici S.p.A. (5).

In contrast to crystallographic fragment screening experiments, compounds in repurposing libraries are chemically more complex (fig. S1A) (6, 7). Thus, these compounds likely bind more specifically and with higher affinity (8). Because of the higher molecular weights, we performed cocystallography experiments at a physiological pH of 7.5 instead of compound soaking into native crystals (9).

From the 5953 compounds in our screen, we obtained x-ray diffraction datasets for 2381 compounds, which we subjected to automated structure refinement followed by cluster analysis (10) and pan dataset density analysis (PanDDA) (11) (table S1). We observed additional density, indicating binding to Mpro, for 43 compounds, which were classified as hits, representing 37 distinct compounds (tables S1, S2, and S3). From these, the binding mode could be unambiguously determined for 29.
molecules (Fig. 1A and table S4). The majority of hits were found in the active site of the enzyme. Of the 16 active site binders, six covalently bind as thiocethers to Cys \textsuperscript{145}, one compound binds covalently as a thioisoximicetial to Cys\textsuperscript{145}, one is zinc-coordinated, and eight bind noncovalently. The remaining 13 compounds bind outside the active site at various locations (Fig. 1A).

Of the 43 hits from our x-ray screen, 37 compounds were available in quantities required for testing their antiviral activity against SARS-CoV-2 in cell assays (table S2). Nine compounds that reduced viral RNA (vRNA) replication by at least two orders of magnitude in Vero E6 cells (fig. S2) were further evaluated to determine the effective concentrations that reduced not only vRNA but also SARS-CoV-2 infectious particles by 50% (EC\textsubscript{50}) (Fig. 2). Additionally, AT7519 and ifenprodil, which showed slightly lower vRNA level reduction, were included because of their distinct binding sites outside of the active site. From these 11, seven compounds (AT7519, calpeptin, ifenprodil, MUT056399, peltinimb, tolylcalpeptin, and triglycidyl isocyanurate) exhibited a >5-fold reduction in infectious particles in combination with either a selectivity index (SI) calculated as the ratio of cytotoxic concentration (CC\textsubscript{50}) divided by the EC\textsubscript{50} of >5 or no cytotoxicity in the tested concentration range and are considered antivirally active (table S5).

Here, we focus on a more detailed description of the 11 compounds analyzed in the secondary screen, which are grouped according to their binding modes to the active site. In both modes, the compound’s central ring sits on top of the catalytic dyad (His\textsuperscript{41}, Cys\textsuperscript{145}), and its three epoxypropyl substituents reach into subsites S\textsubscript{1}, S\textsubscript{1’}, and S\textsubscript{2}. The noncovalent binding mode is stabilized by hydrogen bonds to the main chain of Gly\textsuperscript{143} and Gln\textsuperscript{166} and to the side chain of His\textsuperscript{163}. Tolperisone is used as a skeletal muscle relaxant (15). The x-ray structure suggests that isofoxethylen binds similarly to a fragment to Cys\textsuperscript{145}, forming a thioether (Fig. 3D). Triglycidyl isocyanurate has been tested as an antitumor agent (16).

Calpeptin shows the highest antiviral activity in the screen (EC\textsubscript{50} = 72 nM, CC\textsubscript{50} > 100 μM). It binds covalently via its aldehyde group to Cys\textsuperscript{145}, forming a thioisothiocetial. This peptidomimic inhibitor occupies substrate pockets S\textsubscript{1} to S\textsubscript{3}, similar to the peptidomimetic inhibitors GC-376 (17, 18), calpain inhibitors (19, N3 (2), and the α-ketoamide 13b (2). The peptidomimetic backbone forms hydrogen bonds to the main chain of His\textsuperscript{163} and Glu\textsuperscript{166}, whereas the norleucine side chain maintains van der Waals contacts with the backbone of Phe\textsuperscript{140}, Leu\textsuperscript{144}, and Asn\textsuperscript{145} (Fig. 3E). Calpeptin has known activity against SARS-CoV-2 M\textsuperscript{pro} in enzymatic assays (17). The structure is highly similar to the common protease inhibitor leupeptin (fig. S3A), which served as a positive control in our x-ray screen but was not tested further in antiviral assays. In silico docking experiments also suggested calpeptin as a possible M\textsuperscript{pro} binding molecule (table S7). Calpeptin also inhibits cathepsin L (20), and dual targeting of cathepsin L and M\textsuperscript{pro} is suggested as an attractive path for SARS-CoV-2 inhibition (19).

MUT056399 binds noncovalently to the active site (EC\textsubscript{50} = 38.24 μM, CC\textsubscript{50} > 100 μM). The diphenyl ether core of MUT056399 blocks access to the catalytic site, which consists of Cys\textsuperscript{145} and His\textsuperscript{41}. The terminal carboxamide group occupies pocket S\textsubscript{1} and forms hydrogen bonds to the side chain of His\textsuperscript{163} and the backbone of Phe\textsuperscript{140} (Fig. 3F). The ethyl phenyl group of the molecule reaches deep into pocket S\textsubscript{2}, which is enlarged by a shift of the side chain of Met\textsuperscript{20} out of the substrate binding pocket. MUT056399 was developed as an antibacterial agent against multidrug-resistant Staphylococcus aureus strains (21).

---

**Fig. 1.** The x-ray screening of drug-repurposing libraries reveals compound binding sites distributed across the complete M\textsuperscript{pro} surface. (A) Schematic drawing of M\textsuperscript{pro} dimer structure. Protomer A is shown in white, and protomer B is in red. For clarity, the 29 binding compounds (yellow sticks) are only depicted on one of the two protomers. Catalytic residues His\textsuperscript{41}(H41) and Cys\textsuperscript{145}(C145), the active site, and two allosteric drug binding sites are highlighted. (B) Close-up view of the active site with peptide substrate bound (blue sticks), modeled after SARS-CoV M\textsuperscript{pro} (PDB 2Q6G). The scissile bond is indicated in yellow and with the green arrowhead. Substrate binding pockets S\textsubscript{1}, S\textsubscript{1’}, S\textsubscript{2}, and S\textsubscript{4} are indicated by colored regions.
Quipazine maleate showed moderate antiviral activity (EC50 = 31.64 µM, CC50 > 100 µM). In the x-ray structure, only the maleate counterion is observed covalently bound as a thioether (supplementary text and fig. S3B). Maleate is observed in structures of six other compounds showing no antiviral activity. The observed antiviral activity is thus likely caused by an off-target effect of quipazine.

In general, the enzymatic activity of Mpro critically depends on the dimerization for dimerization (25). The mutation Arg298Ala causes a reorientation of the dimerization domain relative to the catalytic domain, leading to changes in the oxyanion hole and destabilization of the SI pocket by the N terminus. AT7519 was evaluated for treatment of human cancers (26). The potential of allosteric inhibition of Mpro through modulation of Arg298 has been independently demonstrated by mass spectrometry (27).

Our x-ray screen revealed 43 compounds binding to Mpro, with seven compounds showing antiviral activity against SARS-CoV-2. We present structural evidence for interaction of these compounds at active and allosteric sites of Mpro, although we cannot exclude that off-target effects played a role in the antiviral effect in cell culture, in particular for compounds with a low selectivity index. Conversely, an absence of antiviral activity of compounds binding clearly to Mpro in the crystal might be due to rapid metabolism in the cellular environment.
Fig. 3. Covalent and noncovalent binders in the active site of M\textsuperscript{3\textalpha}. Bound compounds are depicted as colored sticks, and the surface of M\textsuperscript{3\textalpha} is shown in gray with selected interacting residues shown as sticks. Substrate binding pockets are colored as in Fig. 1. Hydrogen bonds are depicted by dashed lines. (A) Tolperisone. (B) HEAT. (C) Isofloxthepin. (D) Triglycidyl isocyanurate. (E) Calpeptin. (F) MUT056399.

Fig. 4. Screening hits at allosteric sites of M\textsuperscript{3\textalpha}. (A) Close-up view of the binding site in the dimerization domain (protomer A, gray cartoon representation), close to the active site of the second protomer (protomer B, surface representation) in the native dimer. Residues forming the hydrophobic pocket are indicated. Pelitinib (dark green) binds to the C-terminal \(\alpha\)-helix at Ser\textsuperscript{303} and pushes against Asn\textsuperscript{142} and the \(\beta\)-turn of the pocket S1 of protomer B (resides marked with an asterisk). The inset shows the conformational change of Gin\textsuperscript{256} (gray sticks) compared with the M\textsuperscript{3\textalpha} apo structure (white sticks). (B) RS-102895 (purple), ifenprodil (cyan), PD-168568 (orange), and tofogliflozin (blue) occupy the same binding pocket as pelitinib. (C) AT7519 occupies a deep cleft between the catalytic and dimerization domain of M\textsuperscript{3\textalpha}. (D) Conformational changes in the AT7519-bound M\textsuperscript{3\textalpha} structure (gray) compared with those in the apo structure (white).

Environment. Calpeptin and pelitinib showed strong antiviral activity with low cytotoxicity and are suitable for preclinical evaluation. In any case, all hit compounds are valuable lead structures with potential for further drug development, especially because drug-repurposing libraries offer the advantage of proven bioactivity and cell permeability (28).

The most active compound, calpeptin, binds in the active site similar to other members of the large class of peptide-based inhibitors that bind as thiohemi-acetals or -ketal to M\textsuperscript{3\textalpha} (29). In addition to this peptidomimetic inhibitor, we discovered several nonpeptidic inhibitors. Those compounds binding to the active site of M\textsuperscript{3\textalpha} contained new Michael acceptors based on \(\beta\)-aminoketones (tolperisone and HEAT). These compounds lead to the formation of thioethers and have not been described as produgs for viral proteases. We also identified a noncovalent binder, MUT056399, that blocked the active site. In addition to this common active site inhibition, we identified compounds that inhibit the enzyme through binding at two allosteric sites of M\textsuperscript{3\textalpha}.

The first allosteric site (dimerization domain) is in the direct vicinity of the S1 pocket of the adjacent monomer within the native dimer. The potential for antiviral inhibition through this site is demonstrated by pelitinib. The hydrophobic nature of the residues forming the main pocket is conserved in all human coronavirus M\textsuperscript{3\textalpha} (fig. S8). Consequently, potential drugs targeting this binding site may be effective against other coronaviruses. The potential of the second allosteric site as a druggable target is demonstrated by the observed moderate antiviral activity of AT7519.

REFERENCES AND NOTES

ACKNOWLEDGMENTS

We acknowledge Deutsches Elektronen-Synchrotron (DESY, Hamburg, Germany), a member of the Helmholtz Association HGF, and the Cluster of Excellence "Advanced Imaging of Matter" (C.I.M.), the Hamburg Institute for the Structure of Matter. The Cluster of Excellence Funding: Consortium. We acknowledge financial support from the European Commission (projects ExNet-00002 and InternLabs-00011 "HIER3X"), the Federal Ministry of Education and Research (BMBF) via projects 05K16GU, 05K19GU4, 05K20BI1, 16GW0277, and 03IB040SD, and the Joachim-Herz-Stiftung Hamburg (project 7AWR, 7AWS, 7AWU, 7AWW, 7AXM, 7AXO, 7AY7, 7A1U, 7ABU, 7ADW, 7AF0, 7AGA, 7AHA, 7AK4, 7AKU, 7AMJ, 7ANS, 7AOL, 7AP6, 7APH, 7AQE, 7AQI, 7AR5, 7AR6, 7ARF, 7AVD, 7AWR, 7AWS, 7AWU, 7AX6, 7AXM, 7AXQ, 7AY7, 7BB3, and 7NEV. Code used in this analysis has been previously published (ID). The code for forcing adherence to the Wilson distribution is included in the Vagabond refinement package (https://vagabond.hggin.co.uk/) under a GPLv3 license. Compounds from the Fraunhofer IME Repurposing Collection were obtained from the Fraunhofer Institute for Molecular Biology and Applied Ecology under a material transfer agreement. Compounds from the Safe-in-Drug Library were kindly provided by Dompé Farmaceutici S.p.A. Other materials are available from Se.G. or A.M. upon request. This work is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. To view a copy of this license, visit https://creativecommons.org/licenses/by/4.0/. This license does not apply to figures/photos/artwork or other content included in the article that is credited to a third party; obtain authorization from the rights holder before using such material.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/372/6452/642/suppl/DC1 Materials and Methods
Supplementary Text
Figs. S1 to S9
Tables S1 to S7
References (30–54)
MDAR Reproducibility Checklist
20 November 2020; accepted 29 March 2021
Published online 2 April 2021
10.1126/science.abb7045
A large-scale screen to target SARS-CoV-2

The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) genome is initially expressed as two large polyproteins. Its main protease, Mpro, is essential to yield functional viral proteins, making it a key drug target. Günther et al. used x-ray crystallography to screen more than 5000 compounds that are either approved drugs or drugs in clinical trials. The screen identified 37 compounds that bind to Mpro. High-resolution structures showed that most compounds bind at the active site but also revealed two allosteric sites where binding of a drug causes conformational changes that affect the active site. In cell-based assays, seven compounds had antiviral activity without toxicity. The most potent, calpeptin, binds covalently in the active site, whereas the second most potent, peltinib, binds at an allosteric site.

Science, this issue p. 642