bioRxiv preprint doi: https://doi.org/10.1101/2020.10.29.339317; this version posted October 30, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

COVID Moonshot: Open Science Discovery of SARS-CoV-2 Main Protease Inhibitors by Combining Crowdsourcing, High-Throughput Experiments, Computational Simulations, and Machine Learning

The COVID Moonshot Consortium[1]

[1]Full list of authors: https://tinyurl.com/y3r7redd

***For correspondence:**
john.chodera@choderalab.org (John Chodera); alpha.lee@postera.ai (Alpha Lee); nir.london@weizmann.ac.il (Nir London); frank.von-delft@diamond.ac.uk (Frank von Delft)

Abstract Herein we provide a living summary of the data generated during the COVID Moonshot project focused on the development of SARS-CoV-2 main protease (Mpro) inhibitors. Our approach uniquely combines crowdsourced medicinal chemistry insights with high throughput crystallography, exascale computational chemistry infrastructure for simulations, and machine learning in triaging designs and predicting synthetic routes. This manuscript describes our methodologies leading to both covalent and non-covalent inhibitors displaying protease IC50 values under 150 nM and viral inhibition under 5 uM in multiple different viral replication assays. Furthermore, we provide over 200 crystal structures of fragment-like and lead-like molecules in complex with the main protease. Over 1000 synthesized and ordered compounds are also reported with the corresponding activity in Mpro enzymatic assays using two different experimental setups. The data referenced in this document will be continually updated to reflect the current experimental progress of the COVID Moonshot project, and serves as a citable reference for ensuing publications. All of the generated data is open to other researchers who may find it of use.

Since early 2020, when the coronavirus SARS-CoV-2 was implicated as the cause of the pandemic soon to engulf the world, the need for targeted therapeutic development has been evident due to the dearth of existing antiviral treatments. Realizing this need, researchers worked rapidly to determine the structure of the novel SARS-CoV-2 main protease (MPro) [1], a promising antiviral drug target involved in SARS-CoV-2 viral replication. Other researchers quickly followed with a crystallographic and electrophilic fragment-screen of the structure [2].

Recognizing the necessity for a large-scale, international effort to further develop antiviral compounds, the COVID Moonshot project was created to build on these early results and advance the development of Mpro inhibitors [3]. The project brought together academic and industrial partners from across the world to combine experimental and computational expertise towards this urgent goal.

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.29.339317; this version posted October 30, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

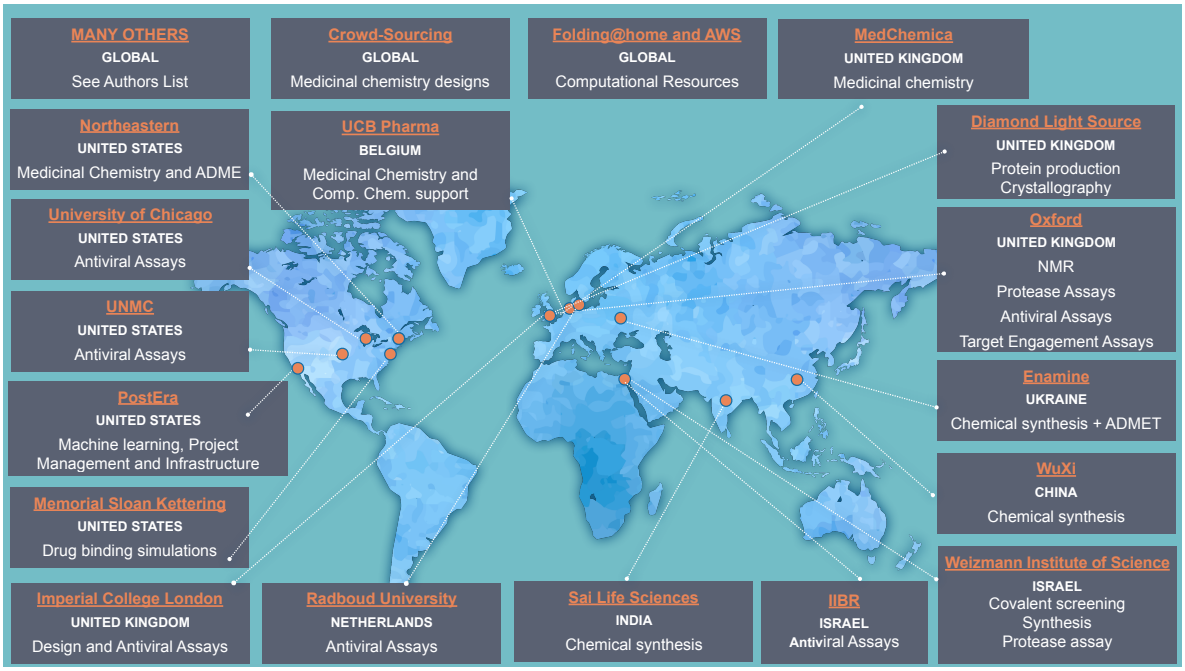

Figure 1. COVID Moonshot is a global open-science effort to develop a patent-free antiviral. The project is an international effort involving over 30 groups (see author list), many of whom have provided work pro-bono or at-cost. The project is IP-free and has relied on philanthropic donations up to this point.

The COVID Moonshot project has focused on progressing early fragment-screening results into potent compounds with activity against both the main protease and the virus. Design of compounds was done through a novel crowd sourcing mechanism, harnessing the expertise of drug-designers across the world. Building on their individual sets of expertise, these designers have entered over 1000 submissions for a total of over 10,000 individual molecule designs to date. The rationales for each design include docking-based approaches, by-eye structure-based designs, machine learning approaches, crawling of the past literature on SARS and MERS compounds, and other general medicinal-chemistry insights that can be visualised at https://postera.ai/covid. In the background, a High-Throughput Screen of 185,000 compounds has been carried out to further aid with lead generation. The results of this screen will be reported in a future manuscript.

Out of these myriad designs, over 1000 compounds have been either ordered or synthesized and tested for activity against MPro in assays at two different sites. The enzymatic activity has been measured using two complementary biochemical assays assessing inhibition of the SARS-CoV-2 main protease: a fluorescence based assay (Weizmann Institute) and RapidFire Mass Spectrometry assay (University of Oxford). Furthermore, structural characterization of selected compounds has resulted in over 200 crystal structures, when combined with the fragment screening data.

Particularly promising compounds have been sent through a further ADMET assay cascade, and sufficiently suitable compounds have been assessed in cellular antiviral screening assays performed by collaborators at the University of Oxford, the University of Nebraska Medical Centre, Israel Institute for Biological Research and the Radboutd University Medical Centre.

All design and experimental data can be found on the COVID Moonshot website at https://postera.ai/covid, which also links to many crystal structures found at https://fragalysis.diamond.ac.uk/viewer/react/preview/target/Mpro. The complete raw data in csv format with detailed description is located at the Moonshot GitHub repository, https://github.com/postera-ai/COVID_moonshot_submissions. The data is frequently updated and we kindly ask that all questions regarding the data be directed at the active COVID Moonshot forum https://discuss.postera.ai/c/covid/general/10.

This is an ongoing project, and this manuscript is very much a "living document". The reason for the publication of this document is to provide readers and ensuing publications a central point to reference

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.29.339317; this version posted October 30, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

data generated from COVID Moonshot.

Figures 2 and 3 exemplify our progress – from crowdsourced fragment merges (TRY-UNI-714a760b-6) to antiviral activity that is comparable to that of Remdesivir.

1 Methodology

1.1 Crowdsourced design campaign

Our approach combines the collective creativity of the scientific community with algorithms that aid with synthesis planning and free energy calculations for prioritizing crowdsourced designs. Specifically, as we started with a dense set of fragment hits with over 70 protein-fragment structures, a crowdsourcing approach allows us to sample different creative merges and expansions of the different fragments.

Those crowdsourced designs are then triaged based on a lax docking-based threshold. Docking was performed by several groups pursuing several different approaches—of both covalent and non-covalent compounds. Consensus docking results were used to filter out designs which demonstrably do not fit the binding site, a difficult task to do by-eye given the sheer number of designs. We then triage compounds based on synthetic accessibility to ensure short cycle times. As the total number of commercially available virtual compounds exceeds 10 billion, keeping track of synthetic accessibility is a non-trivial task. Complex-looking molecules might be synthesizable in a single step, while simple functionalizations could require a complex multi-step synthesis. During the project, consortium member PostEra built a new tool (https://postera.ai/manifold) in order to accelerate this difficult task. Additional custom synthesis tools from PostEra allow for rapid prioritization of compounds and generation of synthetic routes. Furthermore, the algorithms are tailored to the building blocks available at each of our partner CROs, such that we can algorithmically allocate load.

The resulting compounds are synthesized at partner CROs, and all compounds are subjected to: (1) MPro RapidFire mass spectrometry assay, (2) MPro fluorescence assay, (3) High throughput X-ray crystallography. Non-covalent compounds are also analysed by NMR. Early on the use of two complementary protease assays was deemed important due to the rapid development of procedures and the need for verification. Agreement between the label free MS based approaches and fluorophore-based approaches have been reassuring, while discrepancies between assays have led to discovery of false-positives. High throughput crystallography has allowed for rapid generation of design ideas that merge early binding fragments with more recent actives generated during the project.

As soon as the data is cleaned and standardized, the results of all assays are released as on the Moonshot platform (http://postera.ai/covid and https://discuss.postera.ai/c/covid/general/10), ready for the crowd to submit followup compounds.

We received over 13,000 diverse designs from more than 400 contributors. Our data-driven approach enabled the synthesis and testing of over 1,000 compounds in less than 6 months. Some members of the community informally organised themselves into a medicinal chemistry design team, with weekly meetings to go through the data and suggest new designs. Moreover, exascale computing resources donated by Folding@Home enable Moonshot to use Free Energy Perturbation to prioritise compounds.

1.2 Fluorescence MPro inhibition assay

Compounds were seeded into assay-ready plates (Greiner 384 low volume 784900) using an Echo 555 acoustic dispenser, and DMSO was back-filled for a uniform concentration in assay plates (maximum 1%). Screening assays were performed in duplicate at 20 μM and 50 μM. Hits of greater than 50% inhibition at 50 μM were confirmed by dose response assays. Reagents for Mpro assay reagents were dispensed into the assay plate in 10μl volumes for a final volume of 20 μL. Final reaction concentrations were 20 mM HEPES pH 7.3, 1 mM TCEP, 50 mM NaCl, 0.01% Tween-20, 10% glycerol, 5 nM Mpro, 375 nM fluorogenic peptide substrate ([5-FAM]-AVLQSGFR-[Lys(Dabcyl)]-K-amide). Mpro was pre-incubated for 15 minutes at room temperature with compound before addition of substrate. Protease reaction was measured continuously in a BMG Pherastar FS with a 480/520 ex/em filter set. Data analysis was performed with Collaborative Drug Discovery (CDD).

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.29.339317; this version posted October 30, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

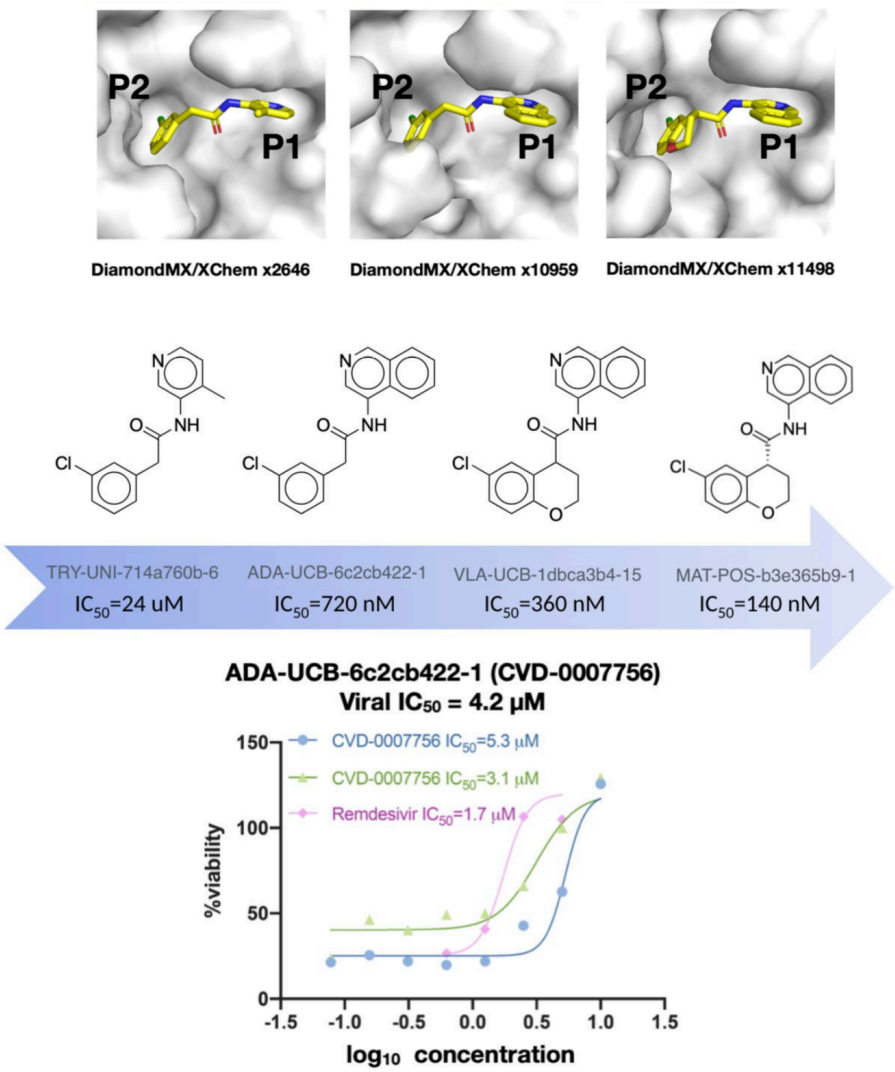

Figure 2. The progression of a fragment merge (TRY-UNI-714a760b-6) into potent, non-covalent MPro inhibitors. One of these compounds ADA-UCB-6c2cb422-1, has also been tested in an antiviral assay and showed promising activity. Antiviral activity was measured at the Israel Institute for Biological Research (IIBR) with Remdesivir serving as the positive control. The more potent analogs of ADA-UCB-6c2cb422-1 have not yet been tested for antiviral activity. The development of these series will be further detailed in a future publication.

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.29.339317; this version posted October 30, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

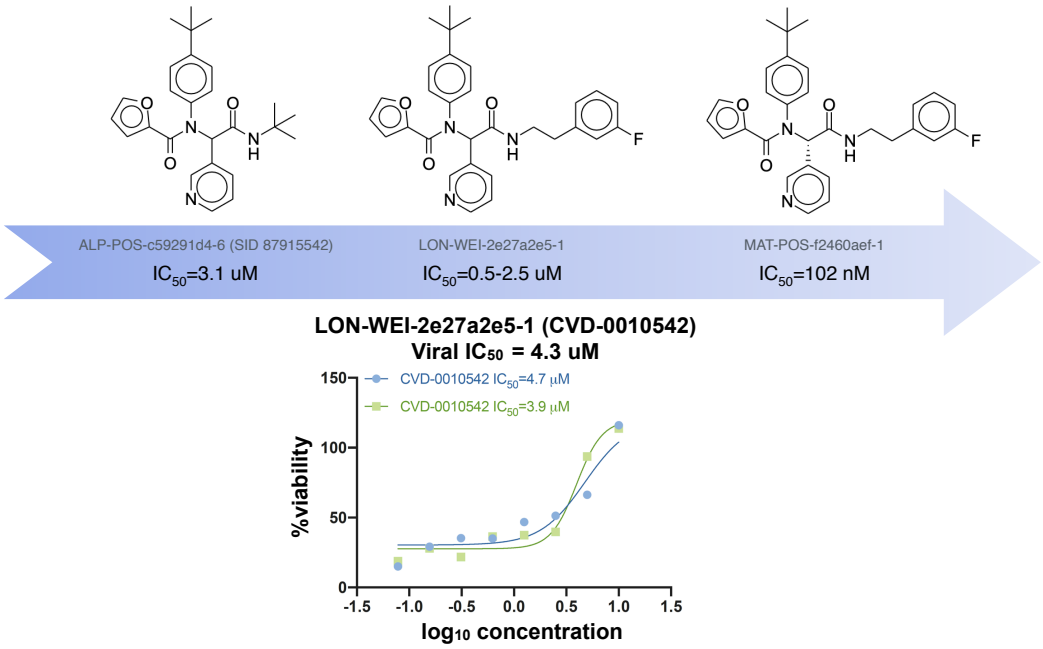

Figure 3. The development of a SARS-CoV-1 non-covalent MPro inhibitor (Jacobs et al. [4]) into a potent SARS-CoV-2 MPro inhibitor with antiviral activity. LON-WEI-2e27a2e5-1 has also been tested in the IIBR antiviral assay and showed activity. The development of these series will be further detailed in a future publication.

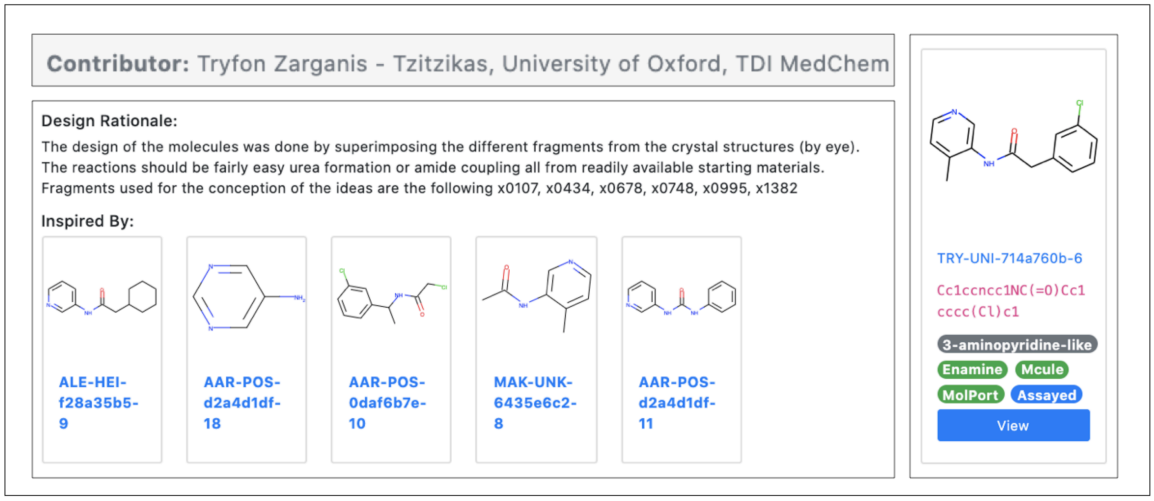

Figure 4. The original crowdsourced design of the fragment merge (TRY-UNI-714a760b-6) that led to the development of the lead series shown in Figure 2. The system to enter compounds with rationales, and add attribution to the original fragment inspirations was built by PostEra specifically for this project. All linked biochemical and structural results on this specific molecule can be viewed on the detail page: https://postera.ai/covid/submissions/714a760b-0e02-4b09-8736-f27f854f8c22/6

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.29.339317; this version posted October 30, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

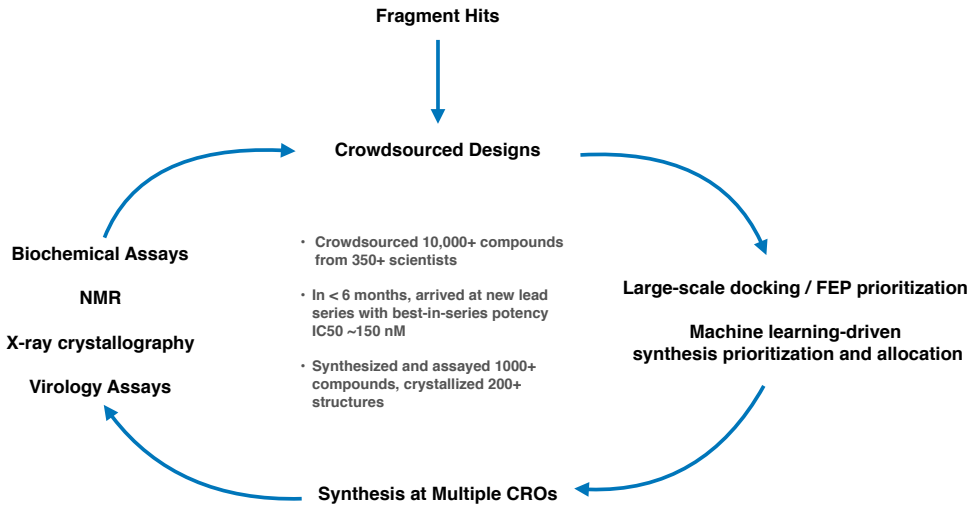

Figure 5. Overall workflow for the COVID Moonshot project. These interconnected steps take place among teams distributed internationally among numerous researchers who have never met in person. Coordination of these synergistic efforts has taken a significant amount of logistical planning amid lockdowns and other complications resulting from the COVID pandemic.

1.3 RapidFire MPro inhibition assay

Inhibitor compounds at 20mM in DMSO are dispensed into 384-well plates (Greiner 384PP 781280) using an ECHO 650T dispenser (DMSO concentration < 1%, final volume = 500 nL.). A 15 µM enzyme stock solution is prepared in 20 mM HEPES, pH 7.5 and 300 mM NaCl, and subsequently diluted to a working solution of 300 nM Mpro in assay buffer (20 mM HEPES, pH 7.5 and 50 mM NaCl) before the addition of 25 µL to each well using a Multidrop Combi (Thermo Scientific). After a quick centrifugation step (1 000 rpm, 15 s) the plate is incubated for 15 min at room temperature. The reaction is initiated with the addition of 25 µL of 4 µM substrate (TSAVLQSGFRK-NH2, initially custom synthesized by the Schofield group, GLBiochem) dissolved in assay buffer. After centrifugation (1000 rpm, 14 s) the reaction is incubated for 10 min at room temperature before quenching with 10% formic acid. The reactions are analysed with MS using RapidFire (RF) 365 high-throughput sampling robot (Agilent) connected to an iFunnel Agilent 6550 accurate mass quadrupole time-of-flight (Q-TOF) mass spectrometer using electrospray. All compounds are triaged by testing the % inhibition at 5 and 50 µM final concentration. Dose response curves uses an 11-point range of 100–0.0017 µM inhibitor concentrations. RapidFire integrator software (Agilent) was used to extract the m/z (+1) charge states of both the substrate (1191.67 Da) and cleaved N-terminal product TSAVLQ (617.34 Da) from the total ion chromatogram data followed by peak integration. Percentage conversion (product peak integral / (product peak integral + substrate peak integral))*100) and percentage inhibitions were calculated and normalised against DMSO control with deduction of any background signal in Microsoft Excel. IC_{50}s were calculated using Levenberg–Marquardt algorithm used to fit a restrained Hill equation to the dose-response data with both GraphPad PRISM and CDD.

1.4 Saturation Transfer Difference NMR (STD-NMR) assay

Method described in Kantsadi and Vakonakis [5].

1.5 High throughput x-ray crystallography

Method described in Douangamath et al. [2].

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.29.339317; this version posted October 30, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

1.6 High Throughput Screening (HTS) of SARS-CoV-2 Mpro

184,631 compounds from the G-INCPM screening collection, as well as a 3,000 Mpro targeted compound library from Enamine, were screened against Mpro using a fluorogenic protease assay. A screening cascade was established as follows: 1) hits were defined as showing a greater than 50% inhibition were selected, filtered for obvious structural alerts and promiscuity in unrelated assays 2) 885 hits were re-tested in three independent Mpro assay runs to confirm activity 3) 128 confirmed hits in 5 concentrations were tested in two independent Mpro assay runs to determine preliminary potency as IC_{50}. Promising compounds were re-sourced, and are awaiting validation.

HTS was performed in both 384-well (20 µL) and 1536-well (10 µM, Greiner 782076) formats. Compounds were pre-plated into barcoded assay ready plates for a final concentration of 10 µM (0.1% DMSO). The assay was performed using Thermo combi dispensers with 1 µL tubing integrated with plate storage and BMG Pherastar FS plate reader and Spinnaker plate mover driven by Momentum software (Thermo). For 1536-well screening, plates were unloaded and spun for 1 min at 1500 RPM prior to plate reading. Data was loaded to Screener software (Genedata) for mapping, normalization and annotation. Curated data was then loaded to CDD Vault for merging with compound annotations.

The results of the HTS screen can be viewed publicly at https://github.com/postera-ai/COVID_moonshot_submissions

1.7 High-throughput solubility threshold measurement

Nephelometry-based solubility assay to define threshold compound solubility at the concentrations of 20 µM and 100 µM in PBS solution. The compound solubilities are normalized to Deoxyfluorouridine (100% soluble) and ondansetron (0% solubility). To put the numbers in perspective, the expected relative solubility values are:

- High >0.8
- Mid 0.6 - 0.8
- Low <0.6

To perform the screening, 245 μL aliquots of PBS buffer were added to each well of 96-well microplates with clear flat bottom. Plates with the buffer were subjected to optical integrity inspection using Nephelostar. Total scanning time for one 96-well plate was 3 min. The pass criterion was set as the background signal in any of the scanned wells below 25 RNU, thus making optical quality of the plates satisfactory for the assay. In case of an excessively high background signal in any wells of the test-plates, those plates/wells were excluded from the study (data not shown). Compounds were obtained from Enamine repository as solids formatted in polypropylene, round bottom blank tubes, in latch boxes (Matrix 4271). Compounds were dissolved in DMSO at 50 mM, incubated at 24–26°C for 8 hours, shaken for 1 hour at 1800 rpm using high-speed microplate shaker Illumina, then incubated at 24–26°C for 14 hours. Intermediate DMSO solutions of the tested compounds were prepared to 2 mM (for the tested concentration 20 µM) and 8 mM (for the tested concentration 100 µM). Ondansetron and DOFU were added to the final plates in 100 mM concentration to get 1 mM and 2 mM final concentrations in the assay plates. To prepare the test plates, 2.5 µL aliquots of DMSO solutions of the tested compounds, reference compounds or pure DMSO were transferred from the polypropylene tubes to the corresponding wells of 96-well plates with PBS buffer using Plate Mate, according to the plate map (Figure 1). Thus, final volumes in the test plates were brought to 247.5 µL, resulting in concentrations of the compounds in test wells of 20 µM and 1% DMSO, correspondingly. Then, turbidity of the solutions was immediately scanned for each well.

1.8 Synthesis prediction

We employ an approach based on the Molecular Transformer technology [6, 7]. Our algorithm uses natural language processing to predict the outcomes of chemical reactions and design retrosynthetic routes starting from commercially available building blocks. This proprietary platform is provided free of charge by PostEra Inc (http://postera.ai). Additionally, Manifold (https://postera.ai/manifold) was built by PostEra Inc. during the

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.29.339317; this version posted October 30, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

project to search the entire space of purchasable molecules, and automatically find the optimal building blocks.

1.9 Viral screening assays

A variety of antiviral replication assays were performed in collaborating laboratories, including cytopathic effect (CPE) inhibition assays at the IIBR, Israel, viral qPCR at Radboud, Netherlands, immunofluorescence assays at University of Nebraska Medical centre, USA, and plaque assays and focus forming unit assays at University of Oxford, UK.

1.9.1 Antiviral assays Cytopathic Effect CPE (IIBR, Ness-Ziona, Israel)

SARS-CoV-2 (GISAID accession EPI_ISL_406862) was kindly provided by Bundeswehr Institute of Microbiology, Munich, Germany. Virus stocks were propagated (4 passages) and tittered on Vero E6 cells. Handling and working with SARS-CoV-2 virus was conducted in a BSL3 facility in accordance with the biosafety guidelines of the Israel Institute for Biological Research (IIBR). Vero E6 were plated in 96-well plates and treated with compounds in medium containing 2% fetal bovine serum. The assay plates containing compound dilutions and cells were incubated for 1 hour at 37°C temperature prior to adding Multiplicity of infection (MOI) 0.01 of viruses. Viruses were added to the entire plate, including virus control wells that did not contain test compound and Remdesivir drug used as positive control. After 72h incubation viral cytopathic effect (CPE) inhibition assay was measured with XTT reagent. Three replicate plates were used.

1.9.2 Antiviral assays qPCR (Radboud University Medical Center, Nijmegen, Netherlands)

Cells. African green monkey Vero E6 kidney cells (ATCC CRL-1586) and Vero FM kidney cells (ATCC CCL-81) were cultured in Dulbecco's modified Eagle medium (DMEM) with 4.5g/L glucose and L-glutamine (Gibco), supplemented with 10% Fetal calf serum (FCS, Sigma Aldrich), 100 µg/ml streptomycin and 100 U/ml penicillin (Gibco). Cells were maintained at 37°C with 5% CO2.

Virus. SARS-CoV-2 (isolate BetaCoV/Munich/BavPat1/2020) was kindly provided by Prof. C. Drosten (Charité-Universitätsmedizin Berlin Institute of Virology, Berlin, Germany) and was initially cultured in Vero E6 cells up to three passages in the laboratory of Prof. Bart Haagmans (Viroscience Department, Erasmus Medical Center, Rotterdam, The Netherlands). Vero FM cells were infected with passage 3 stock at an MOI of 0.01 in infection medium (DMEM containing L-glutamine, 2% FCS, 20 mM HEPES buffer, 100 µg/ml streptomycin and 100 U/ml penicillin). Cell culture supernatant containing virus was harvested at 48 hours post-infection (hpi), centrifuged to remove cellular debris, filtered using a 0.2 μm syringe filter (Whatman), and stored in 100 μl aliquots at -80°C.

Compounds. 10mM compound stocks dissolved in DMSO were provided by Enamine (Ukraine).

Virus stock titration. Vero E6 cells were seeded in 12-well plates at a density of 500,000 cells/well. Cell culture medium was discarded at 24 h post-seeding, cells were washed twice with PBS and infected with 10-fold dilutions of the virus stock. At 1 hpi, cells were washed with PBS and replaced with overlay medium containing Minimum Essential medium (MEM), 2% FCS, 20mM HEPES buffer, 100μg/ml streptomycin, 100 U/ml penicillin and 0.75% carboxymethyl cellulose (Sigma Aldrich). At 48 hpi, the overlay medium was discarded, cells were washed with PBS and stained with 0.25% crystal violet solution containing 4% formaldehyde for 30 minutes. Thereafter, staining solution was discarded, plates washed with double-distilled water, dried and plaques were counted.

Antiviral assay. Vero E6 cells were seeded onto 24-well plates at a density of 150,000 cells/well. At 24 h post-seeding, cell culture medium was discarded, cells were washed twice with PBS and infected with SARS-CoV-2 at an MOI of 0.01 in the presence of six concentrations of the inhibitors (25μM – 0.06μM). At 1 hpi, the inoculum was discarded, cells were washed with PBS, and infection medium containing the same concentration of the inhibitors was added to the wells. SARS-CoV-2 infection in the presence of 0.1% DMSO was used as a negative control. At 24 hpi, 100 μl of the cell culture supernatant was added to RNA-Solv reagent (Omega Bio-Tek) and RNA was isolated and precipitated in the presence of glycogen according to manufacturer's instructions. qRT-PCR. TaqMan Reverse Transcription reagent and random hexamers (Applied Biosystems) were used for cDNA synthesis. Semi-quantitative real-time PCR was performed using GoTaq

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.29.339317; this version posted October 30, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

qPCR (Promega) BRYT Green Dye-based kit using primers targeting the SARS-CoV-2 E protein gene (Forward primer, 5-ACAGGTACGTTAATAGTTAATAGCGT-3'; Reverse primer, 5-ACAGGTACGTTAATAGTTAATAGCGT-3'). A standard curve of a plasmid containing the E gene qPCR amplicon was used to convert Ct values relative genome copy numbers.

Cell viability. Vero E6 cells were seeded in 96-well white-bottom culture plates (Perkin Elmer) at a density of 30,000 cells per well. At 24 h post-seeding, cells were treated with the same concentrations of compounds as used for the antiviral assay. Cells treated with 0.1% DMSO were used as a negative control. At 24 h post-treatment, cell viability was assessed using the Cell Titer Glo 2.0 kit (Promega) using the Victor Multilabel Plate Reader (Perkin Elmer) to measure luminescence signal.

1.9.3 Antiviral assays Immunoflourescence (Pathology and Microbiology, University of Nebraska Medical Centre, USA, St Patrick Reid)

Vero E6 cells were pretreated with 20 uM of the Moonshot compounds for around 2h. Cells were then infected with SARS-CoV-2 at a MOI of 0.1 for 24h. Virus infection was terminated by 4% PFA fixation. Cells were stained using a Rabbit SARS-CoV-2 antibody (Sino Biological40150-R007) as a primary antibody, and Alexa-488, Hoechst and Cell Mask (Thermo Fisher) as a secondary antibody. Images were collected on the Operetta system imaging system, and analysed using the Harmony software.

1.9.4 Antiviral assays Plaque assay and Focus Forming Unit Assay (University of Oxford, UK)

Cell culture. The African green monkey Vero E6 cell line (ATCC CRL-1586) was cultured in Dulbecco's modified Eagle medium (DMEM) with Glutamax supplemented with 100 µg/mL streptomycin, 100 U/mL penicillin, and 10% heat-inactivated fetal calf serum (FCS). The human lung cancer cell line Calu-3 (Anderson Ryan, Department of Oncology, Medical Science Division, University of Oxford) was cultured in a 1:1 mixture of DMEM with Glutamax and Ham's F-12 medium supplemented with 100 µg/mL streptomycin, 100 U/mL penicillin, and 10% heat-inactivated FCS. All cells were maintained as mycoplasma free, with regular verifications by polymerase chain reaction (PCR).

Virus propagation. SARS-CoV-2 England/2/2020 was provided at passage 1 from Public Health England, Collindale. Passage 2 submaster and passage 3 working stocks were produced by infecting Vero E6 cells at a multiplicity of infection of 0.01 in virus propagation medium (DMEM with Glutamax supplemented with 2% FCS) and incubating until cytopathic effect was visible. The cell supernatant was then centrifuged at 500 g for 5 minutes, aliquoted and stored at -80°C. The titre of viral stocks was determined by plaque assay. All subsequent assays were performed using a passage 3 stock.

Cell viability. Cell viability was was measured using the CellTiter 96 R AQueous One Solution Cell Proliferation MTA (3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H - 15 tetrazolium, inner salt) Assay (Promega) according to the manufacturer's instruction after treatment with compound. Briefly, Calu 3 cells were treated with compounds in quadruplicate for 3 days. Wells with 200 µL growth medium with and without cells were included as controls in quadruplicate. After the incubation, 100 µL of growth medium was removed and 20 µL of MTS reagent was added to the remaining medium in each well. After a further one to two hour incubation, the absorbance at 490 nm was measured on a Molecular Devices SpectraMax M5 microplate reader.

Antiviral assays. Calu 3 cells were seeded into 96 well plates (VWR) at a cell density of 8×10^4 cells/well. After 48 hours, the wells were inoculated for 90 minutes with 50 L SARS-CoV-2 England 2/2020 at a M.O.I. of 0.5. The inoculum was replaced with 200 L growth medium containing the appropriate compound dilutions or matched controls and incubated for 3 days. The supernatant was harvested and stored at -80°C prior to analysis by plaque or focus forming unit (FFU) assay. For plaque assays, four 10-fold serial dilutions of each supernatant to be analysed were prepared in virus propagation medium. 100l of each dilution was inoculated in triplicate into wells of a 24-well plate followed by 0.5 mL of Vero E6 cells at 5×10^5 cells/mL in virus propagation medium. The plates were incubated for 2 hours before the addition of 0.5 mL of 1.8% CMC overlay (1:1 mix of 3.6% carboxymethylcellulose in H_2O and virus propagation medium). The plates were incubated for a further 4 days before staining with Napthol blue black. Plaques were counted by eye. For Focus forming unit assays, a SARS-CoV-2 Microneutralization assay from the W James lab (Dunn School

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.29.339317; this version posted October 30, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

of Pathology, University of Oxford) was adapted for use as a FFU assay. Briefly, 3 half log dilutions of each supernatant to be analyzed were prepared in virus propagation medium. 20 µL of each dilution was inoculated into wells of a 96-well plate in quadruplicate followed by 100 L Vero E6 cells at 4.5×10^5 cells/mL in virus propagation medium. The plates were incubated for 2 hours prior to the addition of 100 L of 1.8% CMC overlay. The plates were incubated for a further 24 hours. After 24 hours the overlay was carefully removed and the cells washed once with PBS before fixing with 50 µL of 4% paraformaldehyde, after 30 minutes the paraformaldehyde was removed and replaced with 100 µL of 1% ethanolamine in PBS. The cells were permeabilized by replacing the ethanolamine with 2% Triton X100 in PBS and incubating at 37°C for 30 minutes. The plates were then washed 3 times with wash buffer (0.1% Tween 20 in PBS) inverted and gently tapped onto tissue to dry before the addition of 50 l of EY2A anti-N human mAb (Arthur Huang (Taiwan)/Alain Townsend (Weatherall Institute of Molecular Medicine, University of Oxford)) at 10 pmol in wash buffer. The plates were rocked at room temperature for 1 hour, washed and incubated with 100l of secondary antibody Anti-Human IgG (Fc-specific)-peroxidase-conjugate produced in Goat diluted 1:5000 at room temperature for a further hour. 50 µL of TrueBlue peroxidase substrate was added to the wells and incubated at RT for 10 min on the rocker, after 10 minutes the substrate was removed and the plates washed with ddH_2O for 10 minutes. The H2O was removed and the plates allowed to air dry. The foci were then counted using an ELISPOT classic reader system (AID GmbH).

References

[1] **Jin Z**, Du X, Xu Y, Deng Y, Liu M, Zhao Y, Zhang B, Li X, Zhang L, Peng C, et al. Structure of M pro from SARS-CoV-2 and discovery of its inhibitors. Nature. 2020; p. 1–5.

[2] **Douangamath A**, Fearon D, Gehrtz P, Krojer T, Lukacik P, Owen CD, Resnick E, Strain-Damerell C, Ábrányi-Balogh P, Brandaõ-Neto J, et al. Crystallographic and electrophilic fragment screening of the SARS-CoV-2 main protease. Nature Communications. 2020; 11:5047.

[3] **Chodera J**, Lee AA, London N, von Delft F. Crowdsourcing drug discovery for pandemics. Nature Chemistry. 2020; 12(7):581–581.

[4] **Jacobs J**, Grum-Tokars V, Zhou Y, Turlington M, Saldanha SA, Chase P, Eggler A, Dawson ES, Baez-Santos YM, Tomar S, et al. Discovery, synthesis, and structure-based optimization of a series of N-(tert-butyl)-2-(N-arylamido)-2-(pyridin-3-yl) acetamides (ML188) as potent noncovalent small molecule inhibitors of the severe acute respiratory syndrome coronavirus (SARS-CoV) 3CL protease. Journal of medicinal chemistry. 2013; 56(2):534–546.

[5] **Kantsadi AL**, Vakonakis I. Rapid assessment of ligand binding to the SARS-CoV-2 main protease by saturation transfer difference NMR spectroscopy. bioRxiv. 2020; .

[6] **Lee AA**, Yang Q, Sresht V, Bolgar P, Hou X, Klug-McLeod JL, Butler CR, et al. Molecular transformer unifies reaction prediction and retrosynthesis across pharma chemical space. Chemical Communications. 2019; 55(81):12152–12155.

[7] **Schwaller P**, Laino T, Gaudin T, Bolgar P, Hunter CA, Bekas C, Lee AA. Molecular transformer: A model for uncertainty-calibrated chemical reaction prediction. ACS central science. 2019; 5(9):1572–1583.