Title

Identification of inhibitors of SARS-CoV-2 3CL-Pro enzymatic activity using a small molecule in-vitro repurposing screen

Authors

Maria Kuzikov[1], Elisa Costanzi[2], Jeanette Reinshagen[1], Francesca Esposito[3], Laura Vangeel[4] , Markus Wolf[1], Bernhard Ellinger[1], Carsten Claussen[1], Gerd Geisslinger[1,11], Angela Corona[3], Daniela Iaconis[5], Carmine Talarico[5], Candida Manelfi[5], Rolando Cannalire[6], Giulia Rossetti[7], Jonas Gossen[7], Simone Albani[7], Francesco Musiani[8], Katja Herzog[9], Yang Ye[10], Barbara Giabbai[2], Nicola Demitri[2], Dirk Jochmans[4], Steven De Jonghe[4], Jasper Rymenants[4], Vincenzo Summa[6], Enzo Tramontano[3], Andrea R. Beccari[5], Pieter Leyssen[4], Paola Storici[2], Johan Neyts[4], Philip Gribbon[1] and Andrea Zaliani[1]

Affiliations

1) Fraunhofer Institute for Translational Medicine and Pharmacology (ITMP) and Fraunhofer Cluster of Excellence for Immune mediated diseases (CIMD), Schnackenburgallee 114, 22525 Hamburg, and Theodor Stern Kai 7, 60590 Frankfurt, Germany

2) Elettra-Sincrotrone Trieste S.C.p.A., SS 14 - km 163, 5 in AREA Science Park 34149 Basovizza, Trieste, Italy

3) Dipartimento di Scienze della vita e dell'ambiente, Cittadella Universitaria di Monserrato, SS-554, Monserrato, Cagliari, Italy

4) KU Leuven, Department of Microbiology, Immunology and Transplantation, Rega Institute for Medical Research, Laboratory of Virology and Chemotherapy, Herestraat 49 - box 1043, 3000-Leuven, Belgium

5) Dompé Farmaceutici SpA, via Campo di Pile, 67100 L'Aquila, Italy

6) Department of Pharmacy, University of Naples Federico II, Via D. Montesano, 49, 80131 Naples, Italy

7) Institute of Neuroscience and Medicine (INM-9)/Institute for Advanced Simulation (IAS-5) and Jülich Supercomputing Centre (JSC) Forschungszentrum Jülich D-52425 Jülich, Germany

8) Laboratory of Bioinorganic Chemistry, Department of Pharmacy and Biotechnology, University of Bologna, Bologna, Italy.

9) EU-OPENSCREEN ERIC, Robert-Rössle-Straße 10, 13125 Berlin, Germany

10) University of Chinese Academy of Sciences, Beijing, 100049, China

11) Institute of Clinical Pharmacology, Goethe-University, Theodor Stern Kai 7, 60590 Frankfurt, Germany
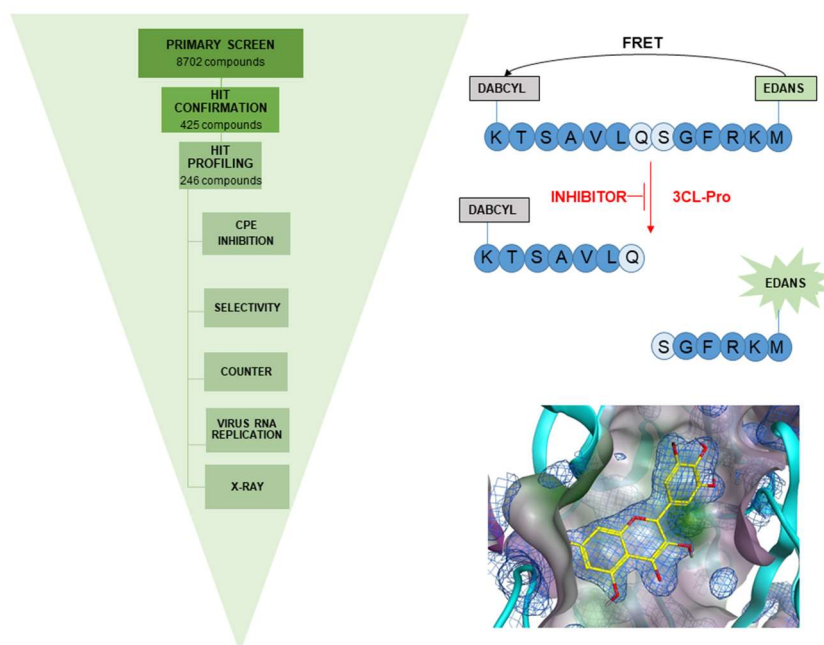
Abstract

Compound repurposing is an important strategy for the identification of effective treatment options against SARS-CoV-2 infection and COVID-19 disease. In this regard, SARS-CoV-2 main protease (3CL-Pro), also termed M-Pro, is an attractive drug target as it plays a central role in viral replication by processing the viral polyproteins pp1a and pp1ab at multiple distinct cleavage sites. We here report the results of a repurposing program involving 8.7 K compounds containing marketed drugs, clinical and preclinical candidates, and small molecules regarded as safe in humans. We confirmed previously reported inhibitors of 3CL-Pro, and have identified 62 additional compounds with IC_{50} values below 1 µM and profiled their selectivity towards Chymotrypsin and 3CL-Pro from the MERS virus. A subset of 8 inhibitors showed anti-cytopathic effect in a Vero-E6 cell line and the compounds thioguanosine and MG-132 were analysed for their predicted binding characteristics to SARS-CoV-2 3CL-Pro. The X-ray crystal structure of the complex of myricetin and SARS-Cov-2 3CL-Pro was solved at a resolution of 1.77 Å, showing that myricetin is covalently bound to the catalytic Cys145 and therefore inhibiting its enzymatic activity.

Keywords
SARS-CoV2 - Main Protease – Screening - FRET – Repurposing

Graphical abstract

Abstract Figure. Workflow for identification and profiling of inhibitors of SARS-CoV-2 3CL-Pro using a large scale repurposing and bioactive compound collection (rhs). Primary assay principle based on quenched FRET peptide substrate of SARS-CoV-2 3CL-Pro (lhs). Inhibiting compounds reduce fluorescence signal relative to DMSO controls. Hit profiling using X-ray.

Introduction

Coronaviruses are enveloped viruses with large (27-31 kb) single-stranded positive-sense RNA genomes. Genes encoding structural and accessory proteins are located at the 3' end which terminates with an UTR and a poly-A tail. The 5' end consists of a 5' cap structure and a non-coding region untranslated region (UTR) followed by open reading frames (ORF) coding for 14-16 non-structural proteins[1]. Past outbreaks of Severe Acute Respiratory Syndrome *coronavirus* (SARS-CoV) and Middle East Respiratory Syndrome (MERS) have demonstrated the zoonotic potential of coronaviruses[2]. The coronavirus SARS-CoV-2 was initially isolated and identified in December 2019 in Wuhan, China and within less than six months developed into a pandemic leading to severe impacts on global health and economic systems. To date (24.11.2020) WHO reports that 64 million people are infected and almost 1.5 million have died[3]. SARS-CoV-2 belongs to the β-coronavirus genus of Coronaviridae family of the order Nidovirales[2]. Non-structural proteins of SARS-CoV-2 take up two thirds of the viral genome and are prominent targets for drugs modulating the replication machinery of the virus[4]. This group includes viral replication and transcription enzymes (RNA-dependent RNA polymerase [RdRp], helicase and 2'O-MTase), the papain-like protease (PLpro), as well as the 3C-like protease (3CL-Pro or M-Pro), the focus of this study.

Role of 3CL-Pro and qualification as an antiviral target

As a single-stranded positive-sense RNA virus, SARS-CoV-2 uses cellular translation machinery directly after infection of the cell to produce viral polyproteins pp1a and pp1ab. The generation of individual viral proteins involves proteolytic cleavage of 3CL-Pro and PLpro, whereby 3CL-Pro is cleaved auto catalytically from the polyprotein, before itself processing the 13 non-structural proteins which go on to initiate further viral replication[1]. Therefore, it plays a central role in progression of infection and generation of new viral particles. Structurally, the SARS-CoV-2 3CL-Pro forms a homodimer, wherein each monomer has three domains[5]. Domains I (residues 8–100) and II (residues 101–183), are located at the N-terminal and are chymotrypsin-like, with six antiparallel beta-barrels forming the substrate binding site located in the cleft between them. Domain III (residues 200–303) at the C-term consists of five helices and is responsible for regulation of dimer formation. Domains I and II are connected to domain III through a loop structure (184 - 199). The active site of SARS-CoV-2 3CL-Pro is formed by a catalytic dyad of Cys-145 and His-41, which is uncommon for Cys protease family enzymes. In common with SARS-CoV 3CL-Pro, a water molecule forms a hydrogen bond with His-41 and can act as a third component in the active site[5]. The SARS-CoV-2 3CL-Pro cleaves at least 11 sites in the polyproteins with preference for Leu-Gln↓(Ser, Ala, Gly) (↓ marks the cleavage position) and a high specificity for the Gln in P1 position, a function only shared with Enteroviruses. Sequence analysis shows that SARS-CoV-2 3CL-Pro shares 96.08 % and 87.00 % sequence identity with the 3CL-Pro of SARS-CoV and MERS, respectively[6]. The similarity of active sites across the three viruses raises the possibility of successful inhibition of SARS-CoV-2 3CL-Pro using compounds previously identified in studies on SARS-CoV and MERS and opportunities for identification of compounds with broad-spectrum activity.

Known inhibitors of 3CL-Pro, examples from SARS-CoV and MERS

Peptidomimetic inhibitors of several viral main proteases have previously been designed according to natural substrate sequences, in which the peptide bonds are chemically modified to

be non-cleavable[7], using designed chemical warheads including carbonyls, hydroxy- and chloromethyl ketones and di-keto moieties. These classes of inhibitors often contain a Michael acceptor moiety and act as reactive substrates. A prominent example is rupintrivir (AG7088), an irreversible inhibitor of the human rhinovirus (HRV) 3C protease, although it lacks activity against SARS-CoV 3CL-Pro, most probably due to the planar fluorophenylalanine at P2[8]. In general, this group of inhibitors exhibit limitations in some relevant medicinal chemistry and pharmacokinetic features such as half-life and stability in human plasma, as well as high binding to plasma proteins which serves to reduce their bioavailability. A second group of 3CL-Pro inhibitors are the alpha-ketoamides peptidomimetics which reversibly inhibit SARS-CoV 3CL-Pro through nucleophilic attachment of the catalytic cysteine to the alpha-keto group. Representatives of this group have been shown to inhibit the MERS 3CL-Pro at sub-micromolar, and SARS-CoV 3CL-Pro at low micromolar, concentrations[5]. A third group is represented by non peptidic reversible inhibitors, among them benzoquinoline compounds, designed using computational modelling. The compounds contain two benzoquinolinones connected to N-phenyl tetrazole moiety through a sulphur atom. The binding mode is proposed to be based on formation of hydrogen bonds and hydrophobic interaction between the S1 and S2 side of 3CL-Pro and benzoquinolinones and N-phenyltetrazole groups, respectively[9]. Pyrazoles have also been shown to inhibit the SARS-CoV 3CL-Pro, as well as 3CL-Pro of type 14 rhinovirus (*RV14*) at low micromolar level, raising the possibility for broad spectrum antiviral compounds to be identified[10].

More recently, in vitro screening of compound libraries and molecular docking based on the crystal structure of the SARS-CoV and MERS 3CL-Pro has revealed inhibitors with potentially different modes of actions against the SARS-CoV-2 3CL-Pro. Among them are metal (zinc)-conjugates (e.g. zinc-pyrithione, PDB: 6YT8)[8,11], natural products of the isoflavone family [12] (e.g. 5,7,3',4'-tetrahydroxy-2'-(3,3-dimethylallyl)-isoflavone), drug candidates (e.g. ebselen) and approved drugs (e.g. disulfiram and carmofur)[6]. To date, virtual screening of large libraries of compounds have been reported (extended list of references and data: https://ghddi-ailab.github.io/Targeting2019-nCoV/computational/) and several in-vitro screens have been published mainly in pre-print form against SARS-CoV-2 3CL-Pro[13].

Compound repurposing

In the case of emerging diseases, such as COVID-19, where limited therapeutic or prevention options in form of vaccine or prophylaxis exist, compound repurposing can allow for fast identification of drug candidates, if the selected therapeutics are efficacious, well-tolerated and have suitable safety profiles. Considerable additional insights into anti-Coronaviridae therapeutic discovery accumulated through past outbreaks of SARS-CoV (2002/2003) and MERS (2012) which informed multiple drug repurposing-based clinical studies following the emergence of SARS-CoV-2. Clinical trials with antivirals (lopinavir, ritonavir) [14] and antimalarials (chloroquine and hydroxychloroquine) [15] have yet to show consistent safety and efficacy. However, other studies including the glucocorticoid, dexamethasone[16], and the viral RNA-polymerase inhibitor remdesivir[17] have led to regulatory approvals for their use in COVID-19 patients. Nevertheless, more effective treatments are still required, in particular for patients at earlier stages of infection, where pharmacological interventions have not been approved so far. Classical antiviral drug development can take many years and need high financial investments. Therefore, repurposing of

known drugs represents a cost and time saving alternative especially when combined with high-throughput screening against relevant disease models or target mechanisms as part of the candidate identification process.

Part of the response to the above need is the project ("EXaSCale smArt pLatform Against paThogEns for CoronaVirus – Exscalate4CoV or E4C", http://www.exscalate4cov.eu) which is funded through EU's H2020-SC1-PHE-CORONAVIRUS-2020 emergency call, (grant N. 101003551). It collects 18 European public Institutions and a private company (DOMPE SpA) in the leading role and brings together high-end computational facilities, drug discovery and BSL-3 high-throughput screening resources. The first aim of E4C is to identify repurposed candidates for a fast track entrance on clinical trials, and in a second phase to provide the foundation for a European platform for rapid response to pandemics with novel drug candidates. In this study, we report on identification of 3CL-Pro inhibitors using repurposed drugs and annotated preclinical compounds, together with a comparative discussion on similar results obtained by cheminformatics and structural biology approaches.

Results and Discussion

Assay development

Enzymatic parameters

Results covering key SARS-CoV-2 3CL-Pro enzymatic properties are shown in Figure 1. Substrate turnover was directly proportional to enzyme concentrations up to 60 nM and assay incubation times up to 15 minutes post substrate addition (Figures 1A and 1B). The relative enzymatic activity was not influenced by the presence of DTT (Figure 1C), with V_{max} = 63070 RFU/min in the presence of 1 mM DTT and 58111 RFU/min without DTT. The corresponding K_m values were 18.8 µM with 1 mM DTT and 16.3 µM without DTT. The assay was not sensitive to DMSO up to a concentration of 5 % v/v (Figure S2).
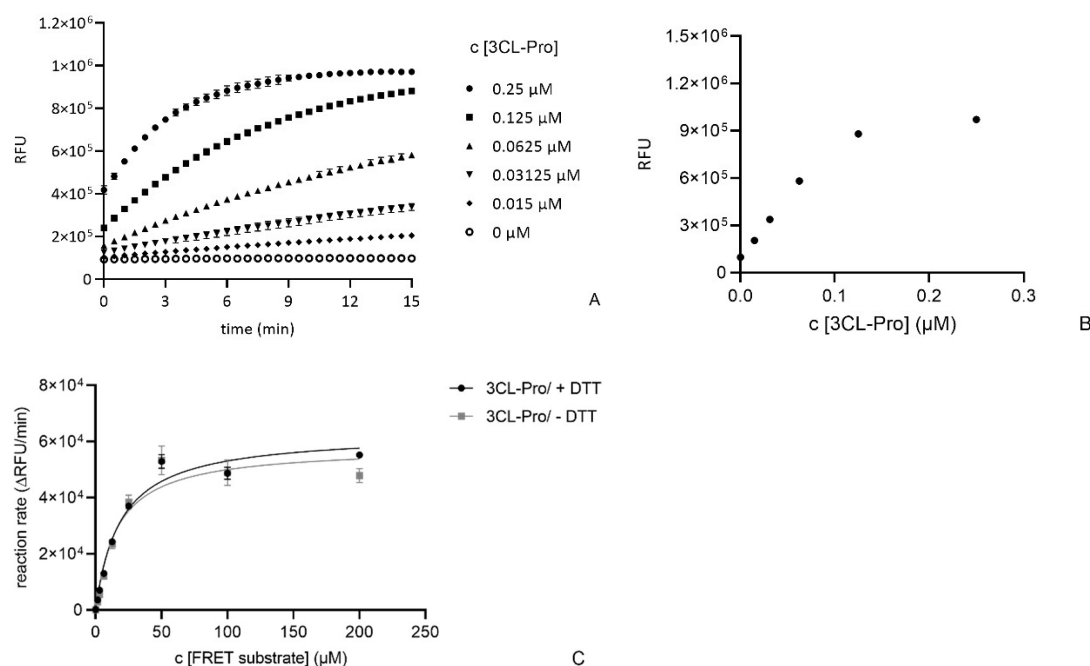
Fig 1. Determination of key enzymatic parameters for SARS-CoV-2 3CL-Pro: A) fluorescence signal versus time for enzyme concentrations from 0 to 0.25 µM (substrate concentration = 15 µM); B) replotting the t = 15 min data against enzyme concentration; C) determination of Km and Vmax in presence and absence of DTT ([SARS-CoV-2 3CL-Pro] = 60 nM). Reaction rate calculations used increase in signal between 0 and 2 minutes.

Compound pre-incubation conditions

Reference inhibitors were calpeptin, a close analogue of GC376, and zinc pyrithione which binds the SARS-COV-2 3CL-Pro catalytic cysteine [8,18]. The effects of time and temperature conditions during compound pre-incubation with SARS-CoV-2 3CL-Pro was determined. Compounds were pre-incubated with the enzyme for 15, 30 and 60 mins (at 25 °C and 20 µM concentration), before substrate addition. The inhibitory effect of zinc pyrithione increased with pre-incubation time, while calpeptin inhibition was unchanged (Figure S3). Similarly, zinc pyrithione was 3-fold more potent after pre-incubation at 37 °C versus pre-incubation at 25 °C (Figure S4). Therefore, to increase assay sensitivity the pre-incubation step in the screening protocol was performed for 60 min at 37 °C.

Effect of DTT

The effect of DTT (0 and 1 mM) on the properties of previously proposed inhibitors of SARS-CoV-2 3CL-Pro was tested at 20 µM compound concentration (Figure S5). Zinc pyrithione, [8] lost its inhibitory effects in the presence of DTT. Chlorhexidine, predicted by docking studies to bind SARS-CoV-2 3CL-Pro[1], was not an effective inhibitor at 20 µM concentration, regardless of DTT. The protease inhibitor, calpeptin[18] retained its inhibitory effect independent of DTT. To account for any DTT dependent effects, primary screening was performed without DTT in the assay buffer.

## Identification of SARS-CoV-2 3CL-Pro inhibitors

The workflow for primary, hit confirmation and profiling stages is shown in Figure 2A.

Primary Screen overview

Primary screen quality control was achieved by analysing the performance of intra-plate controls (Z'-factor) and the reproducibility of duplicate compounds located in the 3 source libraries (Figure S6). The Z' factor values in the primary screen all exceeded 0.68, with an average of 0.86 +/- 0.05, indicating acceptable assay performance (Table S1). The frequency distribution of SARS-CoV-2 3CL-Pro inhibition (%) of the 8702 primary screen compounds is shown in Figure 2B. Individual compound inhibition data and associated annotations are in Table S2 and can be accessed at ChEMBLDB 28 using the AID code CHEMBL4495564[19]. Comparing duplicate test compounds which originated from different sources in the primary screen gave an R^2 = 0.77 (Figure S6), indicating consistency in compound potency between different source libraries (Figure S1). Compounds on each plate were classified as hits if: (test compound inhibition) > (mean + 6x standard deviations of corresponding plate DMSO control inhibition values). A total of 425 compounds with inhibition in primary screening were prioritized for hit confirmation (HC) in the original assay format (in triplicate). Compounds which showed optical interference (elevated fluorescence intensities at 60 min) were removed from the workflow.

Hit confirmation

All Z' values at HC exceeded 0.6, indicating good assay quality. Comparison of Primary and HC inhibition values (Figure 2C) gave R^2 = 0.65. The primary annotated targets of the 8702 screened compounds and the 328 confirmed hits (Hit Rate = 3.8 %) were compared to identify possible enrichment effects (Figure 2D). However, the distribution of target classes associated with the confirmed hit population was similar to that of the screened set.
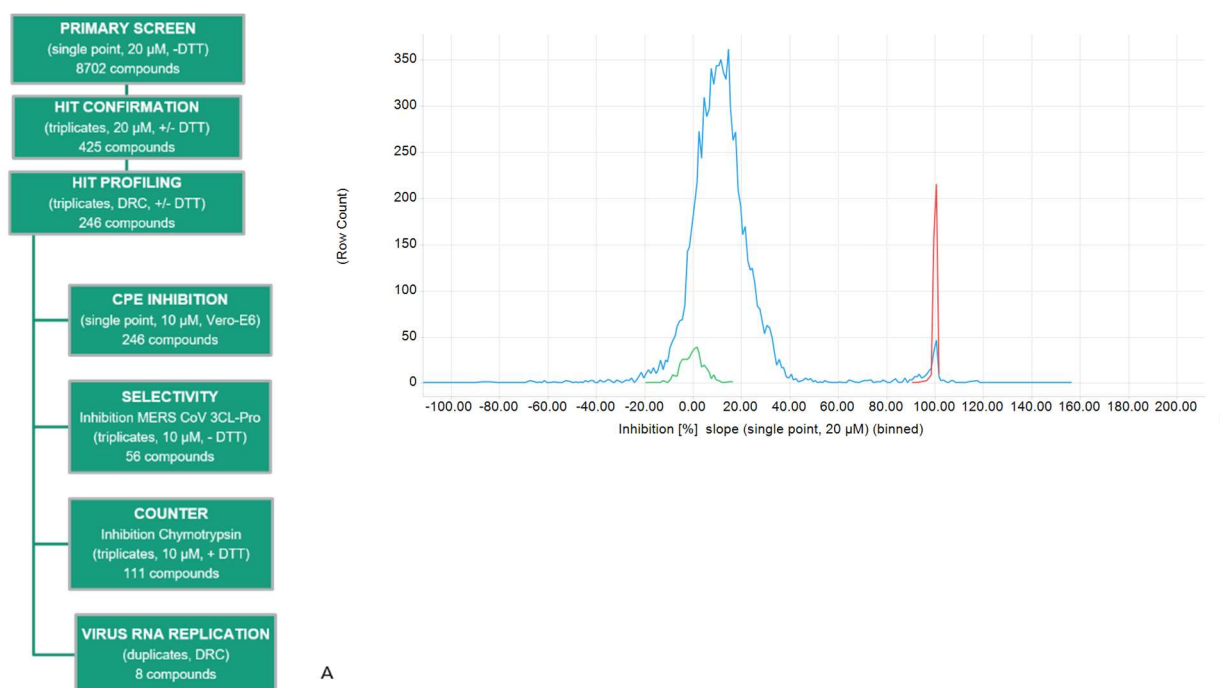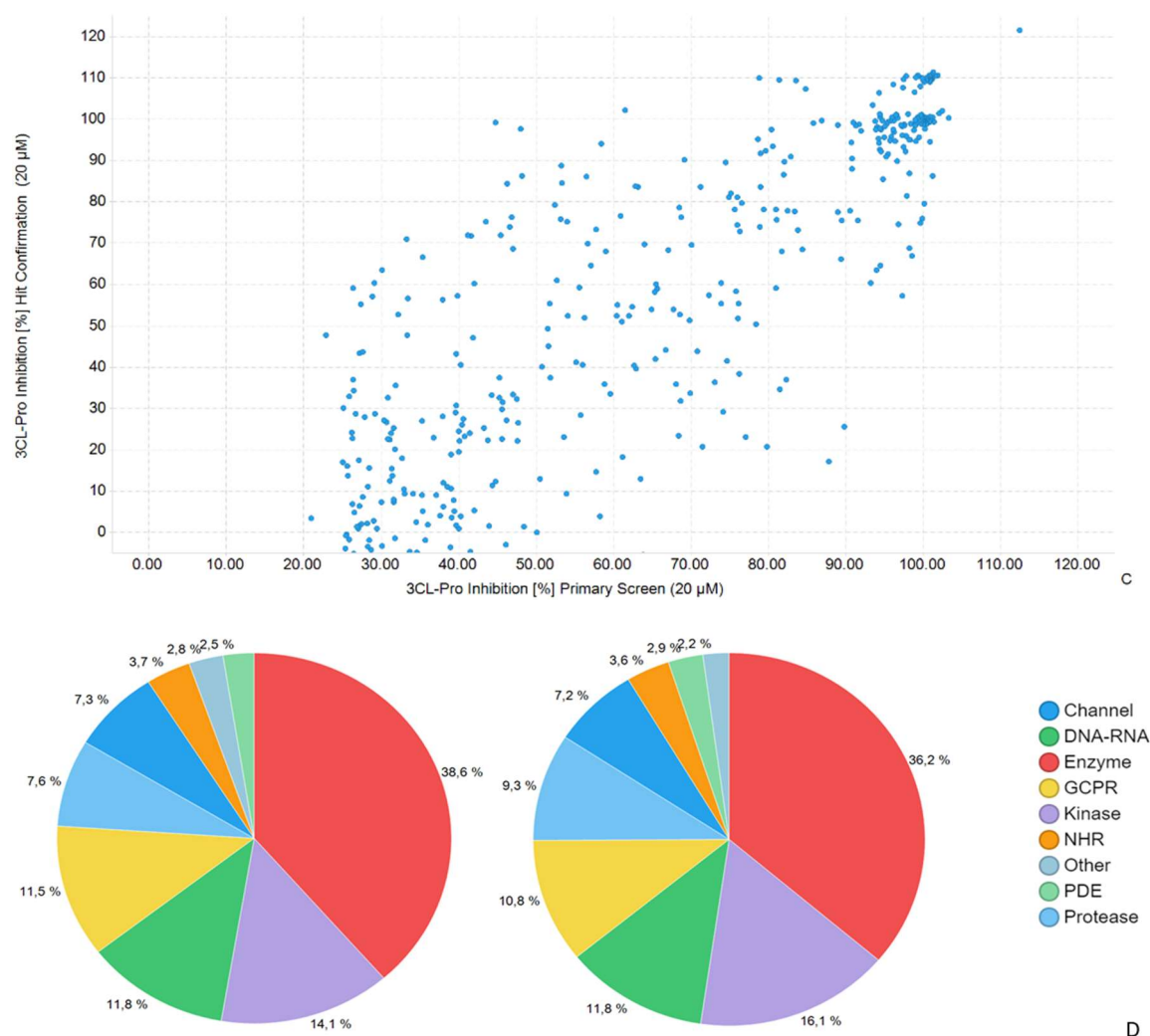
Fig 2. Overview of primary and hit confirmation screening: 2A) Screening workflow; 2B) Primary screen, frequency distribution of inhibition for test compounds and controls; 2C) Inhibition at primary versus hit confirmation for 425 selected inhibitors, R^2=0.65; 2D) compound with annotated primary target classes for screened set (LHS) and 328 confirmed hit compounds (RHS).

Hit profiling overview

A set of 246 compounds with mean SARS-CoV-2 3CL-Pro inhibition exceeding 50 % at HC were then progressed to Hit Profiling (HP). Firstly, the effects of DTT on compound inhibition levels were assessed at 20 µM test compound. Some 156 compounds showed a relative reduction in SARS-CoV-2 3CL-Pro inhibition of more than 30 % in the presence of 1 mM DTT, and these were designated as "DTT sensitive". The remaining 89 compounds gave less than 30 % change in inhibition in the presence of 1 mM DTT and these were designated as "DTT insensitive".

All 246 compounds were profiled in dose response in the absence of DTT, and the 89 DTT insensitive compounds were also tested in the presence of DTT (Table S3). Without DTT, some 119 of 246 compounds gave IC_{50} values below 5 µM for SARS-CoV-2 3CL-Pro inhibition. In the presence of 1 mM DTT, 40 compounds gave an IC_{50} below 5 µM for SARS-CoV-2 3CL-Pro inhibition (Table S3).

A subset of mainly DTT insensitive compounds were then analysed for selectivity towards Chymotrypsin, and the MERS 3CL-Pro (Table S3). Assay buffers for Chymotrypsin contained 1mM DTT. In general, this group of compounds were not active against Chymotrypsin, with only 7 compounds showing inhibition > 30 %. Seven compounds showed inhibitory effect on MERS 3CL-Pro with IC_{50} values < 10 µM, among them PX-12 (IC_{50} = 7.6 µM) for which an IC_{50} of 21 µM has been reported[6]. In the Vero-E6 phenotypic assay only 22 compounds from the 246 selected SARS-CoV-2 3CL-Pro inhibitors showed > 10 % inhibition of CPE at 10 µM. In further studies on 2 unique compounds (MG-132 and thioguanosine) gave an IC_{50} < 20 µM in CPE inhibition assay. These compounds have also been shown to active against Caco2 cells in a recent study[20]. Interestingly, MG-132 was more potent in the Vero-E6 CPE assay (IC_{50} = 0.36 µM), and viral replication assays (EC_{50} = 0.12 µM) (Figure S7), that against SARS-CoV-2 3CL-Pro (IC_{50} = 13.1 µM).

Compounds previously reported as inhibitors of SARS-CoV-2 3CL-Pro

The screen identified 6 compounds (TDZD-8, carmofur, tideglusib, ebselen, disulfiram and baicalein) with IC_{50} values < 1µM, and which have been previously reported as inhibitors of SARS-CoV-2 3CL-Pro (Figure 3 and Table S3). In general, these compounds were 10 fold more potent in our hands compared to literature reports, which may be due to the extended pre-incubation of compounds with SARS-CoV-2 3CL-Pro for 60 minutes at 37 °C. The compounds TDZD-8 and carmofur were insensitive to DTT, whilst the other 4 compounds in this group lost activity in the presence of 1 mM DTT (Figure 3, Table S3). In the selectivity screening, carmofur was also weakly active against Chymotrypsin (14 % inhibition).
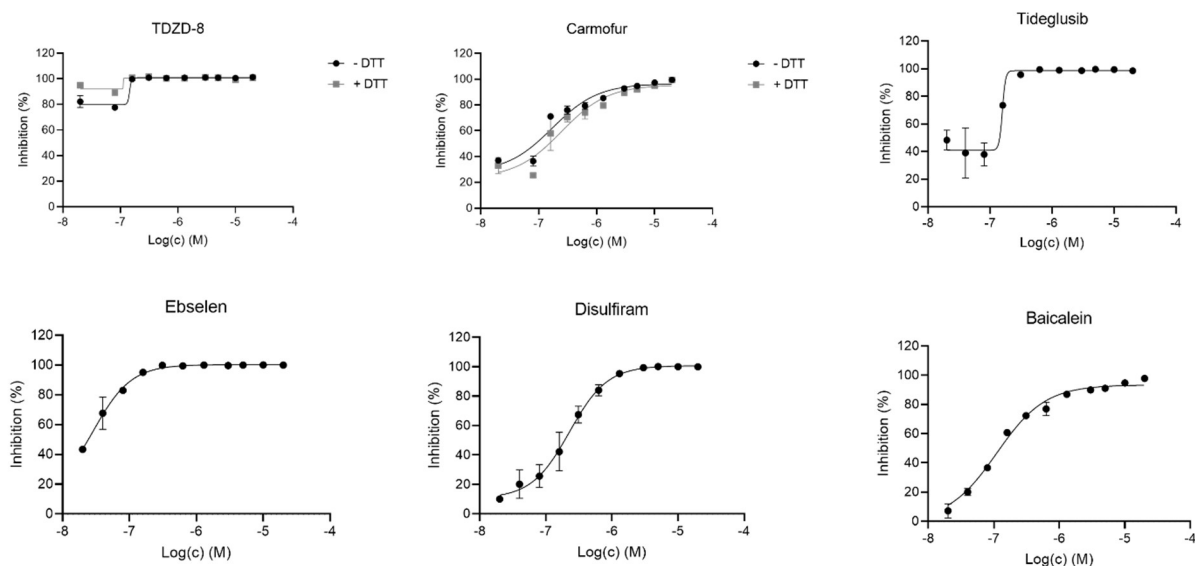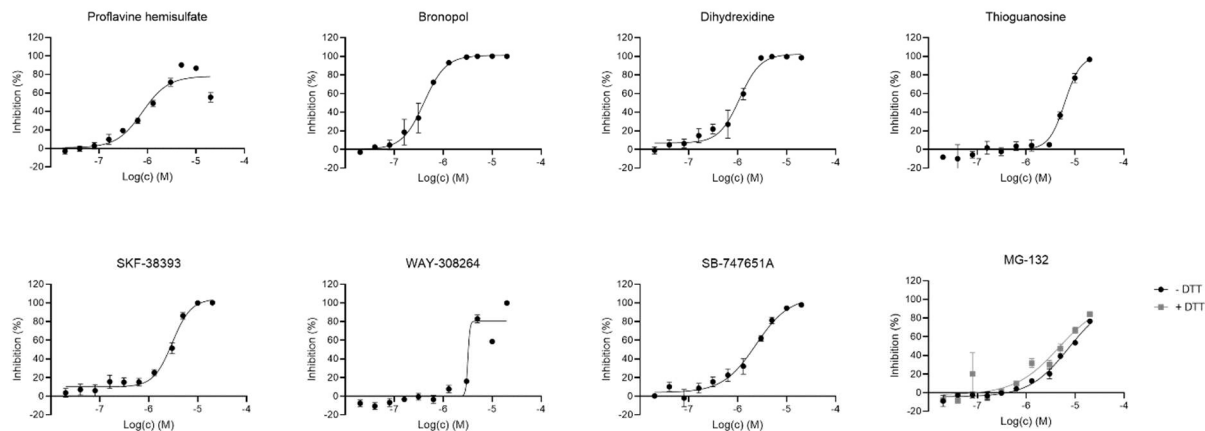

NAME	Target	Phase	PS 3CL-Pro Inhibition [%] (20 µM)	HC 3CL-Pro Inhibition [%] (20 µM)	HP 3CL-Pro IC50 (µM)	HP 3CL-Pro IC50 + DTT (µM)	MERS 3CL-Pro inhibition [%] (10 µM)	IC50 MERS 3CL-Pro (µM)	Chymotrypsin Inhibition [%] +DTT (10 µM)	Cytopathic effect Vero-E6 inhibition [%] (10 µM)	Literature SARS-CoV-2 3CL-Pro IC50 (µM)
TDZD-8	GSK3b inhibitor	Preclinical	100.92	109.50	0.14	0.14	n.t.	n.t.	4.03	2.59	2.15 *
Carmofur	Antimetabolite	Withdrawn	101.09	110.56	0.16	0.35	n.t.	n.t.	14.02	-2.58	1.82 *
Tideglusib	GSK3b inhibitor	Preclinical	100.84	99.48	0.08	n.t.	22.00	n.t.	30.69	-6.7	1.55 *
Ebselen	Glutathion reductase mimetic	Preclinical	99.99	100.38	0.03	n.t.	14.00	n.t.	n.t.	-2.87	0.67 *
Disulfiram	Aldehyde dehydrogenase inh.	Approved	99.73	99.92	0.22	n.t.	12.00	26.75	n.t.	-21.56	9.35 *
Baicalein	Prostaglandin inhibitor	Phase 2	97.8	95.96	0.02	n.t.	47.00	n.t.	n.t.	5.07	0.95 **

Fig.3. Hit profiling for known SARS-CoV-2 3CL-Pro inhibitors identified in the primary screen, n.t. = not tested. (Only DTT insensitive compounds were tested in the presence of DTT).

In our hands, these compounds did not show activity in the CPE Vero-E6 CPE phenotypic assay and they were also inactive in a recently reported study which evaluated SARS-CoV-2 anti CPE phenotypes in a Caco2 cell line[20]. Interestingly, epigallocatechin gallate, a compound structurally related to baicalein, was also active against SARS-CoV-2 3CL-Pro but did not show DTT dependence (IC_{50} without DTT = 1.58 µM, with DTT 0.42 µM), (Table S3). * = ref 5, ** = ref 21

SARS-CoV-2 3CL-Pro inhibitors with moderate anti-CPE effect

The second group of confirmed hits is formed by compounds not previously reported as inhibitors of SARS-CoV-2 3CL-Pro but which showed moderate inhibition (between 10 and 19 %) of CPE in Vero-E6 cells (Figure 4 and Table S3). With the exception of MG-132, inhibition of SARS-CoV-2 3CL-Pro for this group was not observed in the presence of DTT (Table S3). Of the 8 compounds, also only thioguanosine and MG-132 produced dose responses in the Vero-E6 CPE assay giving IC_{50} values of 3.97 and 0.36 µM, respectively. Corresponding cytotoxicity CC_{50} values in Vero-E6 imaging assay were > 20 µM and 2.9 µM, respectively. MG-132 showed a similar reduction in vRNA, with an IC_{50} value of 0.12 µM (Figure S7). Thioguanosine has previously been shown to have anti-SARS-CoV-2 CPE effects in a Caco-2 cell based assay[21].
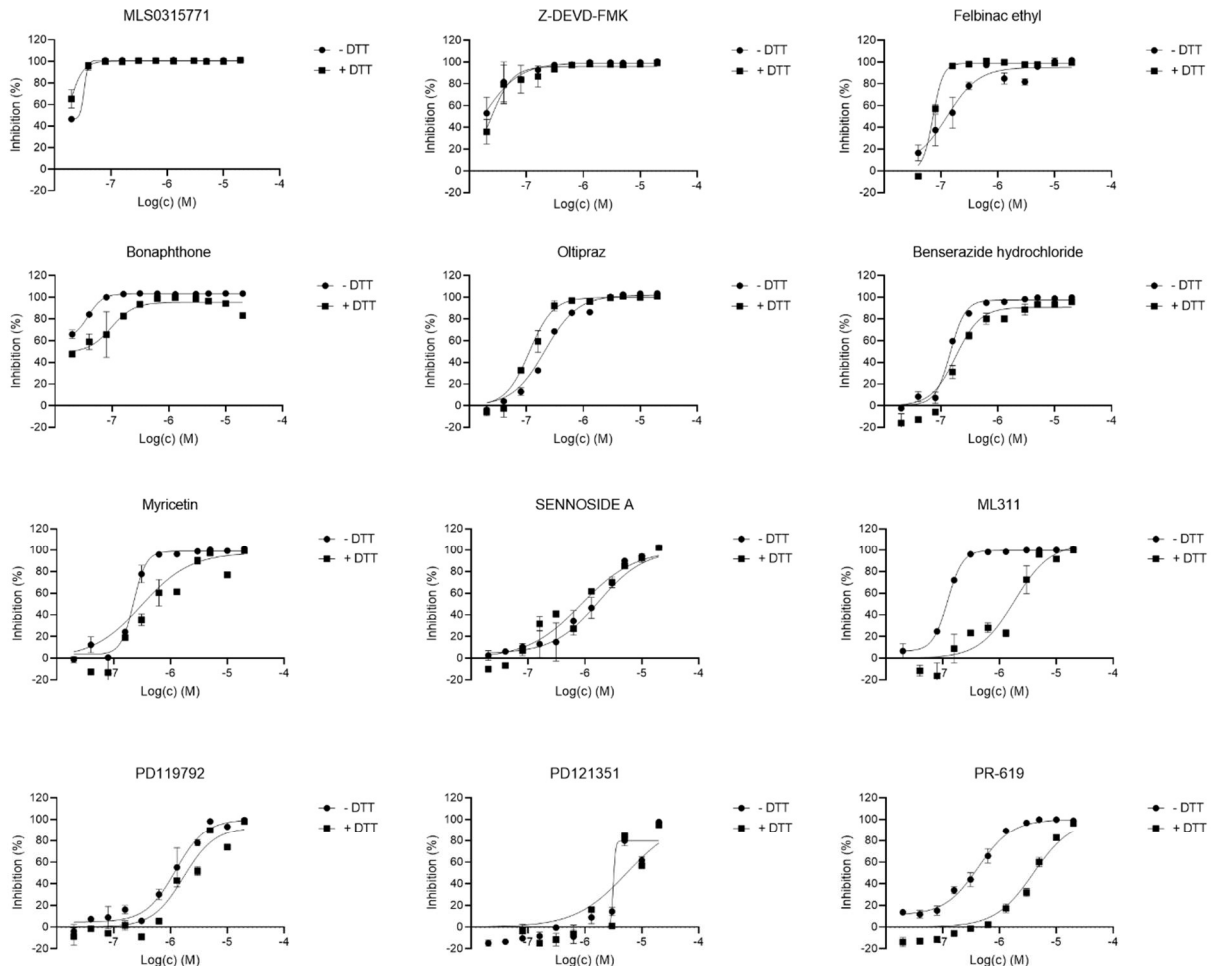

NAME	Target	Phase	PS 3CL-Pro Inhibition [%] (20 µM)	HC 3CL-Pro Inhibition [%] (20 µM)	HP 3CL-Pro IC50 (µM)	HP 3CL-Pro IC50 + DTT (µM)	vRNA reduction EC50 Vero-E6 cells (µM)	MERS 3CL-Pro inhibition [%] (10 µM)	Chymotrypsin Inhibition [%] +DTT (10 µM)	Cytopathic effect Vero-E6 inhibition [%] (10 µM)	CPE IC50 VERO-E6 (µM)
Proflavine hemisulfate	Intercalator	Approved	60.89	76.55	0.83	n.t.	n.t.	37.00	n.t.	19.09	n.t.
Bronopol	Food biocide	Phase 3	99.89	100.48	0.42	n.t.	n.t.	14.00	n.t.	17.6	n.t.
Dihydrexidine	dopamine agonist	Preclinical	98.33	98.95	1.05	n.t.	n.t.	19.00	n.t.	16.67	n.t.
Thioguanosine	Antimetabolite	Phase 2	100.08	79.58	6.31	n.t.	n.a.	n.t.	n.t.	16.05	3.97
SKF-38393	dopamine agonist	Preclinical	99.85	99.87	2.63	n.t.	n.t.	5.00	n.t.	14.07	n.t.
WAY-308264		Preclinical	89.41	66.18	3.17	n.t.	n.t.	n.t.	2.86	11.42	n.t.
MG-132	proteasome inhibitor	Preclinical	68.52	78.61	7.35	4.74	0.12	n.t.	2.52	-12.18	0.36

Fig.4. Dose dependent effects of SARS-CoV-2 3CL-Pro inhibitors which also demonstrated a moderate anti-cytopathic effect in Vero-E6 cells. n.t. = not tested. n.a. = not active.

Compounds inhibiting SARS-CoV-2 3CL-Pro without anti-CPE effects

The third group of compounds act as inhibitors of SARS-CoV-2 3CL-Pro, both in the presence and absence of DTT (IC_{50} < 5 µM in both conditions). Here, we highlight compounds which are selective also against chymotrypsin (inhibition < 10 % at 10 µM) and exhibit promising medicinal chemistry properties (Figure 5 and Table S3). The group contains mainly preclinical compounds which are not ready for immediate repurposing, with the exception being the anti-Parkinson Disease compound benserazide.

NAME	Target	Phase	PS 3CL-Pro Inhibition [%] (20 µM)	HC 3CL-Pro Inhibition [%] (20 µM)	HP 3CL-Pro IC50 (µM)	HP 3CL-Pro IC50 + DTT (µM)	vRNA reduction EC50 Vero-E6 cells (µM)	MERS 3CL-Pro inhibition [%] (10 µM)	IC50 MERS 3CL-Pro (µM)	Chymotrypsin Inhibition [%] +DTT (10 µM)	Cytopathic effect Vero-E6 inhibition [%] (10 µM)
MLS0315771	PMI inhibitor	Preclinical	100.74	110.52	0.02	0.01	n.t.	n.t.	n.t.	-0.52	3.57
Z-DEVD-FMK	Caspase 3 inhibitor	Preclinical	99.95	109.58	0.01	0.01	n.t.	n.t.	n.t.	11.21	7.51
Felbinac ethyl	PPARg inhibitor	Preclinical	96.2	99.77	0.20	0.08	n.a.	18.00	n.t.	-0.88	-5.74
Bonaphthone	Antiviral	Preclinical	101.99	101.38	0.04	0.09	n.a.	9.00	n.t.	8.47	-9.91
Oltipraz	NQO1 inhibitor	Phase 3	101.16	100.56	0.21	0.11	n.a.	10.00	n.t.	9.89	1.98
Benserazide hdrochloride	DOPA decarb. inhibitor	Phase 3	94.29	106.43	0.14	0.18	n.t.	n.t.	n.t.	0.39	-12.61
Myricetin	JAK1 inhibitor	Preclinical	97.44	109.73	0.22	0.26	n.t.	n.t.	n.t.	3.78	-3.28
Sennoside A	Aquaporin inhibitor	Preclinical	103.3	100.33	1.59	0.65	n.t.	42.00	n.t.	2.66	3.34
ML311	MCL1-BIM inhibitor	Preclinical	100.32	110.09	0.15	1.34	n.t.	n.t.	n.t.	3.72	-10.46
PD119792		Preclinical	100.32	109.52	1.23	1.56	n.t.	n.t.	n.t.	9.67	-10.29
PD120339		Preclinical	95.97	100.57	>20	>20	n.t.	n.t.	n.t.	9.35	-3.36
PR-619	Girdin inhibitor	Preclinical	97.57	98.84	0.41	4.41	6.60	100.00	2.25	4.59	-34.29

Fig. 5. Dose dependent effects of 3CL-Pro inhibitors which demonstrated inhibitory effect independent from presence of DTT in the assay media. n.t. = not tested

Myricetin:MPro X-ray Crystallography results

The X-ray crystal structure of the complex of myricetin and SARS-Cov-2 3CL-Pro was solved at a resolution of 1.77Å (Figure 6, lhs and PDB 7B3E). The 2Fo-Fc map unambiguously shows that myricetin is covalently bound to the catalytic Cys145, with a bond length of 1.7 Å. The orientation found is determined in large part by the covalent bond between Cys145 sulphur and the 2' position of the flavonoid leading to unprecedented binding for a flavonoid scaffold. Consistent with the measured complex structure, the non-covalent in-silico docking calculations for myricetin docking led to a similar pose (Figure 6, rhs). Moreover, the 2Fo-Fc map of the myricetin:MPro X-ray structure showed that the binding pocket is only partially occupied by myricetin, and voids are filled by solvent (ethylene glycol and water) molecules, signalling an opportunity for future structure-based drug design efforts.
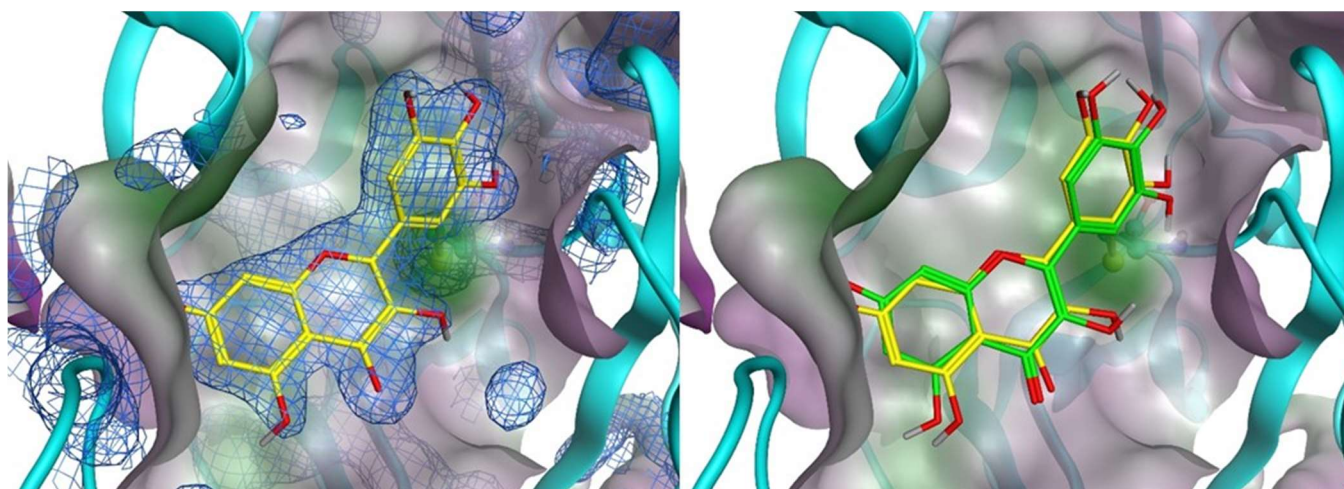

Fig. 6 Xray crystal structure of myricetin covalently bound to the catalytic Cys145 of 3CL-Pro (yellow) with 2Fo-Fc electron density map contoured at 1 sigma (blue grid) on left side (PDB: 7B3E). Cys 145 sidechain is highlighted with ball-stick format, while compounds are depicted with lines. On right side the overlap of the experimental X-ray complexed myricetin (yellow) with the non-covalent docking pose of myricetin (green).

Discussion

The protease 3CL-Pro is responsible for accurately processing the SARS CoV-2 polyprotein at 11 distinct cleavage sites. With the unique exception of the specificity for Gln in P1, the high number of different cleavage sites recognized suggests a certain degree of promiscuity with regard to cleavage site selectivity on other sub pockets. This is consistent both with the diversity of peptide-like compounds found as inhibitors so far and with recent comparison of crystal structures between the 3CL-Pro apo-form obtained at room temperature with the corresponding low-temperature apo- and inhibitor bound-forms, which show large changes in catalytic cavity dimensions upon ligand binding[22]. Molecular dynamics simulations also predict high flexibility in solution, especially of the loop regions adjacent to binding sites[23]. The protein is suggested to work as a dimer with N- and C-terminal regions of each monomer actively involved in ligand binding. The 3CL-Pro catalytic cavity includes four exposed cysteines, of which Cys145 appears predisposed to reacting with covalent ligands. This differs from other Cys-Proteases where a triad (Cys-His-Asp/Glu) is typically involved[24]. In addition, there are also four exposed histidine residues

which may be responsible for other 3CL-Pro biochemical characteristics observed in this study, including DTT sensitivity. These structural properties, combined with the absence of homologous human proteases, point towards anti-3CL-Pro compounds, alone or in combination with other therapies, as having the capacity to exert antiviral effects and thus be a potential treatment of COVID-19 infection.

Reference compounds

In this study, zinc pyrithione, an antifungal and bacteriostatic compound was used as positive control (IC_{50} = 0.05 µM) and was independently identified as a screening hit. This compound was previously shown to inhibit SARS-CoV 3CL-Pro (IC_{50} = 0.17 µM)[8,11] by binding to its catalytic cavity [25] (PDB: 6YT8), suggesting a similar interaction mode for this compound with SARS-CoV-2 3CL-Pro. We reconfirmed the activity of published inhibitors including: TDZD-8, carmofur, tideglusib, ebselen, disulfiram and baicalein which gave IC_{50}'s between 20 and 220 nM. Analogues of ebselen (IC_{50} = 30nM) carrying the same N-phenyl-1,2-benzothiazol-3-ones scaffold, also stood out in the confirmed hit population and the preclinical compounds PD08290 (IC_{50} = 20 nM), MLS0315771 (IC_{50} = 20 nM), PBIT (IC_{50} = 20 nM) and ML345 (IC_{50} = 20 nM) showed higher potencies than ebselen itself (Table S3). The 60 min preincubation time at 37°C facilitated the identification of slowly binding putative cysteine-reactive inhibitors. This type of inhibitory mechanism is supported by the loss of inhibitory capacity for ebselen, disulfiram, tideglusib and baicalein in the presence of reducing conditions (1 mM DTT). Moreover, we confirmed the Caspase 3 inhibitor, Z-DEVD-FMK, as active (IC_{50} = 10 nM) with 600-fold higher potency than observed previously[13]. However, this compound did not influence CPE in phenotypic assays.

Compounds with anti-cytopathic effects

A general correlation between 3CL-Pro and CPE inhibition in the Vero-E6 model was not found (Table S3). Of the anti- CPE compounds, only thioguanosine and MG-132 produced a measurable dose response in the CPE assay (IC_{50} values of 3.97 and 0.36 µM, respectively). Previously, thioguanosine, has been reported to be an inhibitor of the PL-Pro's of SARS-COV[26] and PEDV[27] (PDB: 5XU8) and to be anti-cytopathic against SARS-CoV-2[28] . Structural evidence of catalytic Cys145 sulphur involvement in thioguanine inhibition of PL-Pro[29] may be mechanistically similar to the effects it exerts in the SARS-CoV-2 3CL-Pro cavity.

The lack of anti-CPE effect for the great majority of identified SARS-CoV-2 3CL-Pro inhibitors may be due to modest cell membrane permeability, poor metabolic stability or efflux effect from Vero-E6 cells which express PgP. A second possible reason might reside on the chemical reactivity of SARS-CoV-2 3CL-Pro inhibitors within cells. For example, the antifungal phenylmercuric acetate (IC_{50} = 10 nM) or anti-arthritis auranofin (IC_{50} = 210 nM) both contain chemical moieties prone to react with nucleophiles such as mercury, zinc or gold[11]. Other compounds have a Michael's acceptor, either already present into the structure (bonaphthone with IC_{50} = 40 nM) or generated in-situ such as the case of the molecule HEAT[25] (PDB: 6YNQ). The observation that several compounds with similar structures showed measurable IC_{50}'s against SARS CoV-2 3CL-Pro in a biochemical assay readouts, however, without influencing the CPE, cast doubts on scaffold selectivity and constitutes a warning for medicinal chemistry lead design.

Compounds inhibiting SARS-CoV-2 3CL-Pro

This screen identified PR-619 to be active against 3CL-Pro from SARS-CoV-2 (IC_{50}= 0.41 µM) and MERS (IC_{50}=2.25 µM). This compound contains two isothiocyanate groups and is a known inhibitor of several deubiquitinating enzymes with IC_{50}s ranging 2-10 µM[30,31]. PR-619 is inactive in cellular phenotypic assay either in Vero-E6 cells and Caco-2[20], suggesting its potential as a lead compound against SARS-CoV-2 may be limited. Among peptide-like molecules calpeptin (IC_{50} = 4 µM) and MG-132 (IC_{50} = 7.35 µM) were identified, while the associated derivative boceprevir gave below 50 % inhibition in HC. These molecules are predicted to act by binding Cys145 in the catalytic domain. Nevertheless, calpeptin and MG-132 should be regarded as reversible inhibitors, as both contain aldehyde warhead groups. The recently discovered CDK degraders CR8-(R)[32] was moderately potent against SARS-CoV-2 3CL-Pro (IC_{50} = 0.84 µM), but with no anti CPE effect. However, its activity on CDK machinery could influence the cell phenotype challenged by viral infection and hence there might be a safety concern for further development of CDK inhibitors. Among natural product based compounds, isatin derivatives have been identified as potential inhibitors[33] , however this was not be confirmed in this study, notwithstanding that the isatin-like "1,2-dicarbonyl" moiety is present in the antiviral bonaphthone (IC_{50} = 40 nM). Recent interest as an anti-influenza agent has been raised for this molecule, although no further development has been recorded[34]. Interestingly, the hypotheses that flavonoids could act as privileged scaffold for both SARS CoV and SARS CoV-2 3CL-Pro ligands[35,12] was confirmed in our investigations. In addition to baicalein, a series of flavonoids with IC_{50} ranging 180 nM to 3.6 µM (Table S3) were identified, and within this group myricetin derivatives were well represented.

The presence of polyhydroxy phenolic moiety seems to provide some interaction advantages, as suggested by the reported SARS CoV-2 3CL-Pro crystal structure with baicalein (PDB: 6M2N). The related dopamine decarboxylase inhibitor benserazide (IC_{50} = 0.14 µM), IGF-1 inhibitor NT-157 (IC_{50} = 0.48), nootropic agent exifone (IC_{50} = 4.38 µM) have in common an identical tri-hydroxy phenolic moiety. Moreover, a similar ortho di-hydroxy phenol moiety is present in the dye haematoxylin (IC_{50} = 0.22 µM), dopamine receptor agonists apomorphine (IC_{50} = 0.52 µM), dihydrexidine (IC_{50} = 1.05 µM) and SKF-38393 (IC_{50} = 2.63 µM) and lipoxygenase inhibitor nordihydroguaiaretic acid (IC_{50} = 2.59 µM). Unfortunately, all these compounds showed DTT sensitivity, likely due to the high oxidation propensity of polyhydroxy phenols[36]. For instance, two polyhydroxy phenol-containing compounds, myricetin and benserazide, have been successfully predicted to have nanomolar affinities based on structural modelling (J. Goosen et al. *submitted*). Polyhydroxy phenolic moieties are, however, considered promiscuous and are found as frequent hitters in large scale assessments of screening library performance[37]. This may not only be due to their redox features but also to the presence of a high number of closely arranged promiscuous hydrogen-bond acceptor/donor sites, which raises the number of protein 3D-pharmacophores satisfiable by such molecular structures. Notwithstanding, these "promiscuity" features, curcumin, benserazide, apomorphine, flavopiridol, anthracycline-based antibiotics and other products have reached approval for clinical use across several indications.

To visualize how hit compounds may bind to the protein, we exemplarily docked MG-132 and thioguanosine, the two molecules shown to inhibit viral CPE, into the main catalytic cavity of 3CL-Pro, as well as myricetin for which confirmatory structural information was obtained (Figure 7).

While for MG-132 the docking pose is straightforward, as many peptide-like structures are oriented in a similar direction, for the smaller and more compact thioguanosine the optimal pose was not as evident. The observation that corresponding compound guanosine did not exert the same inhibition potency in the screen, prompted us to focus on the sulphur atom. Through extensive DFT calculations we assessed that, in contrast to guanosine, the most probable tautomer for thioguanosine involves a thiol group rather than a thio-ketonic group (Supplementary modelling section). When docked this tautomer clearly shows its sulphur atom in proximity to the catalytic Cys145, allowing the ribose moiety to make hydrogen-bond interactions with Glu166 backbone atoms (Figure 7). When in such high proximity, thiol-redox reactions to generate S-S species are possible, not only for thioguanosine but for the other potent sulphur-containing compounds which were identified (Table S3). Myricetin appears bound to 3CL-Pro in an opposite orientation to that previously seen with baicalein [21] (PDB: 6M2N). The unprecedented sulphur addition to 2' position of myricetin suggest the presence of a quinoid species apt to be attacked by a strong nucleophile with consequent oxidative loss of an electron or a hydride. This event could be supported by the lack of reducing agents in the crystallization medium like DTT**Error! Bookmark not defined.**.

Only 20 % of the 66 compounds identified with IC_{50} <1 µM had molecular weights above 400 Da. The average molecular weight (MW) of the screened set was 364.9±143.4, with 3114 having MW > 400 (35.8 %). The hypothesized 3CL-Pro cavity flexibility could therefore reward smaller compounds compared to larger ones, either because the protease induced-fit requires more time to successfully complete, or because catalytic Cys145 needs more time to address adjacent electrophiles. In this direction structural differences in cavity binding between nanomolar and micromolar affinity ligands have been illuminating [J. Goosen et al. *submitted*]. We postulated therefore that SARS CoV-2 3CL-Pro analysed in our biochemical assay setup may reward smaller ligands.
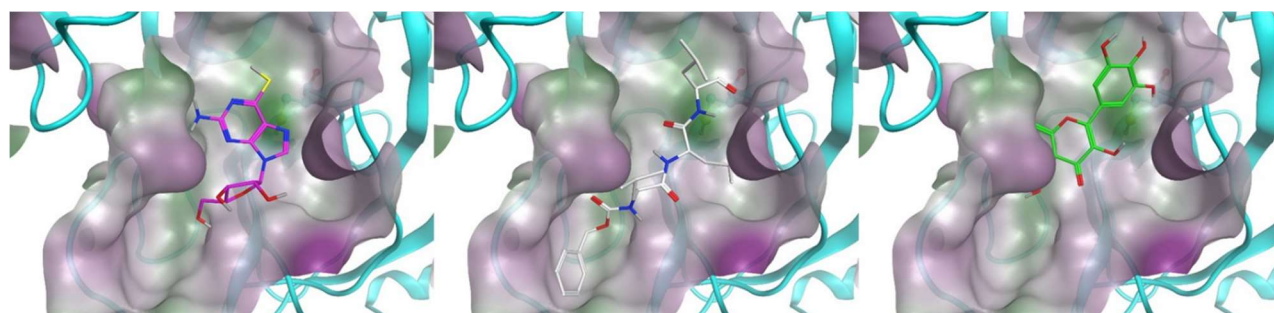

Fig 7 Thioguanosine (magenta), MG-132 (white) and myricetin (green) docked into monomer 3CL-Pro. Catalytic Cys145 are highlighted in ball and stick, while ligands are in stick representation. Protein cavity water-accessible surface is coloured by hydrophilicity. Most probable thioguanosine tautomer with thiol form has been considered according to QM calculations (see supplementary materials).

The sub-micromolar inhibition level achieved by myricetin (IC50=0.22uM) against SARS-CoV-2 3CL-Pro, however, was not reflected in anti CPE or inhibition of viral replication in Vero-E6 cells

(Table S3). Even though the similar compound baicalein did show anti-cytopathic effect in Vero-E6 cells and relative RNA reduction in the cell after 24 h incubation[21]. The longer cellular readout we used (72hrs), might allow metabolic and/or efflux processes to play a role in diminishing the internal molecular concentration of myricetin. However, the information collected with this protein complex can be valuable to drive future structure-based design attempt for flavonoid scaffolds.

Promising candidates for new lead generation

This study identified several novel potent inhibitors of SARS CoV-2 3CL-Pro including calpeptin, CR8-(R), bonaphthone, myricetin, MLS0315771, ML311 and thioguanosine, which show potential for further investigation. Among these compounds there are known scaffolds including peptide derivatives, flavonoids and kinase and protease inhibitors. The diverse structural characteristics of identified actives can provide new opportunities to medicinal chemists as part of further compound property optimization. A promising candidate for further studies is the peptide-like proteasome inhibitor MG-132 (IC_{50} = 7.35 µM), which notwithstanding a low potency on 3CL-Pro, did show anti CPE cytopathic effect on Vero-E6 cells (IC_{50} = 0.36 µM) and confirmed potency on viral RNA replication with an IC_{50} = 0.12 µM by qPCR (Figure S7). The compound is known to be active against Coronaviruses[38,39] and further in-vivo studies would be valuable.

Repurposed candidates

Focusing on already commercialized actives and more promising candidates for clinical developments, the aromatic-L-amino-acid decarboxylase inhibitor benserazide, antiviral riodoxol, dopamine agonist and phosphodiesterase inhibitor apomorphine and proton pump inhibitor rabeprazole may be interesting to study in more detail. All these molecules can be administered orally with measured acute toxicity in rodents higher than 5 mg/kg dose. Benserazide and apomorphine, for instance, can assure exposure in brain and riodoxol has been already been used as an antiseptic and disinfectant, in ointments in the treatment of chronic skin diseases and eczema of a subacute character[40] while rabeprazole has been used in the clinic as antibacterial against Helicobacter Pylori. In-vitro assays determining their limited effects on viral CPE and replication may not fully reflect the infection pathology. We would suggest that additional investigation in other screening formats using, for example, human derived lung tissue samples may establish if the efficacy observed against the SARS-CoV-2 3CL-Pro can be reflected in anti-viral phenotypes for these compounds.

Materials and Methods

Library composition

The screened compounds were sourced from 3 collections. Firstly, the Dompe "Safe-In-Man" (SIM) proprietary collection containing 607 drug candidates which have undergone at least successful Phase I studies. Secondly, the EU-OPENSCREEN collection of 2463 compounds annotated in line with the Drugs & Probes database (https://www.probes-drugs.org/compoundsets), which samples drugs and drug candidates in several development

phases together, along with preclinical probes with high affinity for their primary targets. Finally, the Fraunhofer repurposing collection, assembled based on the design feature the Broad Repurposing collection[41] (Drug Repurposing Hub https://www.broadinstitute.org/drug-repurposing-hub). The Fraunhofer Repurposing Library contains 5632 compounds including 3,400 compounds that have reached clinical use across 600 indications, as well as 1582 preclinical compounds with varying degrees of validation. Overall, 8702 compounds were available for screening. The compound collections overlapped in terms of identity as shown in Figure S1. This overlap was useful in determining the consistency in compound response for material sourced from different collections (Fig. S1). In general, all test compounds were quality controlled by LC/MS for purity and identity (Purity > 90 %) and were stored at -20 °C in 100 % DMSO prior to use.

Source of proteins

The 3CL-Pro of SARS-CoV-2 was synthesised using the ORF1ab polyprotein residues 3264-3569, (GenBank code: MN908947.3). Gene synthesis, protein production and purification was as reported by Zhang et al.[5], where eluted fractions containing the target protein were pooled and subjected to buffer exchange in 20 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA,1 mM DTT, pH 7.8. The 3CL-Pro MERS gene (NCBI reference Sequence: NC_019843.3) was purchased from GenScript. The 3CL-Pro MERS protein was also expressed as described in Zhang et al.[5] Briefly, the 3CL-Pro MERS expression pellets were clarified by ultracentrifugation and purified using a Ni-Sepharose column and by HiTrap Q HP column. Eluted fractions containing the target protein were pooled and subjected to buffer exchange in 20 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, 1 mM DTT and pH 7.8. Protein purity was confirmed by SDS-PAGE analysis.

Primary Assay development

The detection of enzymatic activity of the SARS-CoV-2 3CL-Pro was performed under conditions similar to those reported by Zhang et al[5]. Enzymatic activity was measured by a Förster resonance energy transfer (FRET), using the dual-labelled substrate, DABCYL-KTSAVLQ↓SGFRKM-EDANS (Bachem #4045664) containing a protease specific cleavage site after the Gln. In the intact peptide, EDANS fluorescence is quenched by the DABCYL group. Following enzymatic cleavage, generation of the fluorescent product was monitored (Ex/Em= 340/460 nm), (EnVision, Perkin Elmer). The assay buffer contained 20 mM Tris (pH 7.3), 100 mM NaCl and 1 mM EDTA. The assay was established in an automated screening format, (384 well black microplates, Corning, #3820) and optimized with respect to assay volume, enzyme concentration, substrate concentration, incubation time and temperature, DMSO tolerance, response to inhibition with known compounds as zinc pyrithione [8] and the effects of reducing agents (DTT).

Primary screen, hit confirmation and profiling

In the primary screen, test compounds (stock at 10 mM in 100 % DMSO), positive (zinc pyrithione (medchemexpress, #HY-B0572) 10 mM in 100 % DMSO) and negative (100 % DMSO) controls, were transferred to 384-well assay microplates by acoustic dispensing (Echo, Labcyte). Plate locations were: test compounds at 20 µM final (columns 1 to 22); positive control zinc pyrithione

at 10 µM final (column 23); and negative control 0.2 % v/v (column 24). 5 µl of SARS-CoV-2 3CL-Pro stock (120 nM) in assay buffer were added to compound plates and incubated for 60 min at 37 °C. After addition of 5 µl substrate (30 µM in assay buffer), the final concentrations were: 15 µM substrate; 60 nM SARS-CoV-2 3CL-Pro; 20 µM compound; and 0.2 % DMSO in a total volume of 10 µL/well. The fluorescence signal was then measured at 15 min and inhibition (%) calculated relative to controls (Envision, PerkinElmer). Results were normalized to the 100 % (positive control) and 0 % (negative control) inhibition. To flag possible optical interference effects, primary assay plates were also read 60 min after substrate addition, when the reaction was complete.

For Hit Confirmation (HC), the putative hits from the primary screen were picked and re-tested in the same primary assay format in triplicate at 20 µM compound. Confirmed compounds were then profiled in triplicate in 11 point concentration responses, starting from 20 µM top concentration with 1:2 dilution steps. Whilst the primary screening assay did not contain DTT, the HC and profiling steps were performed +/- DTT (@ 1 mM) to identify any DTT-dependent effects.

Selectivity assays (MERS 3CL-Pro and Chymotrypsin)

A MERS 3CL-Pro inhibition assay was performed to observe potential broad inhibition effects of compounds against related Coronaviruses. The substrate was identical to that used in the SARS-CoV-2 3CL-Pro primary assay. The assay buffer contained 20 mM Tris (pH 7.3), 100 mM NaCl and 1mM EDTA. Test compounds were pre-incubated with 2.1 µM of enzyme (2 µl) and 0.5% DMSO, for 60 minutes at 37 °C prior to addition of substrate to obtain a final concentration of 15 µM substrate and 20 µM test compound in a 20 µl total volume. The signal was monitored after 30 minutes of incubation. A dose response curve of the compound GC376 was used as a positive control[42] and DMSO (0.5 % final) was used as a negative control. Although, showing low similarity to human proteases and being a cysteine protease, SARS-CoV-2 3CL-Pro is structurally related to the serine protease of the chymotrypsin family with a cysteine replacing the serine in the active site [8,43]. Therefore, a Chymotrypsin activity assay kit (#K352-100, Biovision) was used according to manufacturer's instructions to profile the relative selectivity of confirmed Hits. Here, the assay buffer contained 1mM DTT, which was essential to observe chymotrypsin enzymatic activity. Cleaved substrate was detected at Ex/Em= 380/460 nm. The positive control (100 % inhibition) was nafamostat (final concentration 5 µM), and the negative control was DMSO at 0.5 %.

Anti-cytopathic effect and virus yield reduction assays

The CPE assay used the African green monkey kidney cell line (Vero-E6) which had previously been engineered to constitutively express GFP[44]. Cells were maintained in Dulbecco's modified Eagle's medium (DMEM; Gibco) supplemented with 10 % foetal calf serum (FCS; Biowest), 0.075 % Sodium Bicarbonate (7.5 % solution, Gibco) and 1x Pen-strep (Gibco) and kept under 5 % CO2 at 37 °C. Assay medium contained 2 % FCS. SARS-CoV-2 strain BetaCov/Belgium/GHB-03021/2020 recovered from a nasopharyngeal swab taken from an asymptomatic patient returning from Wuhan, China in the beginning of February 2020 was sequenced on a MinION platform (Oxford Nanopore). After serial passaging on Huh7 and Vero-E6 cells, infectious content of the virus stock was determined by titration on Vero-E6 cells using the Spearman-Kärber method. To measure

inhibition of the SARS-CoV-2 cytopathic effect, 384-well plates (Greiner #781092) were spotted with test compounds using an acoustic dispenser (Echo, Labcyte) to yield 10 µM final concentration at 0.1 % DMSO in primary screening. The day before infection (Day -1), plates were equilibrated to room temperature and 30 µL of Vero-E6 EGFP cells were added to give 8,000 cells/well. On the day of infection (Day 0), plates were transported to the CAPS-IT robotic system for the addition of virus dilution (MOI = 0.001) using a liquid handler (EVO 100, Tecan) to a final well volume of 60 µl, and left for incubation at 37 °C, 5 % CO2. At four days post-infection, GFP signal was captured using wide field fluorescence imaging by exciting at 485-20 nm and emitting with the BGRFRN filter set. A 5 X objective captured 80 % of an entire well on a 384 plate at once. The optimal exposure time was determined based on fluorescence intensity and was set on 0.023 seconds. A 2x2 binning was used and autofocus plane count was reduced to increase image acquisition speed. An image analysis protocol was developed in-house by using the SpotDetector bioapplication (Cellomics, Thermofisher). After background reduction on the raw image files, a fixed fluorescent intensity threshold was determined for the identification of GFP cells. Afterwards, the total area ('total amount of surface covered by fluorescent cells') was calculated for the detected cells and compared to the positive (remdesivir 20 µM) and negative (DMSO) control. This parameter directly correlates to cell confluence. Compound profiling used the screening primary protocol as above, with compounds arrayed in 8 point concentration responses in triplicate (max 20 µM, 1:3 dilution steps). Assessment of the underlying cytotoxicity of the compounds was performed as described above in dose response but without virus infection and using Sodium-Selenite (20 µM final) as the cytotoxicity control.

For selected compounds, efficacy was verified in a virus yield reduction assay. Briefly, VeroE6 cells were seeded in a 96-well cell culture plate in a density of 40,000 cells per well. After 24 hours, compounds were added to the medium, starting at a highest concentration of 25 µM and then serially diluted 1:3 (8 steps). Each compound was carried out in duplicate. After two hours, cells were inoculated with virus (MOI = 0.1) and left for incubation. Two hours later, the virus-containing medium was washed away with PBS and replaced with fresh medium and compound. At two days post infection, cells were visually inspected for cytopathic effects and supernatant was collected in the lysis buffer. RNA extraction was carried out using the NucleoSpin kit (Macherey-Nagel) according to manufacturer's instructions. Subsequent RT-qPCR was performed on a LightCycler96 platform (Roche) based on SARS-CoV-2 N gene RNA amplification using forward (5'-GACCCCAAAATCAGCGAAAT) and reverse (5'-TCTGGTTACTGCCAGTTGAATCTG) primers and probe (5'- FAM-ACCCCGCATTACGTTTGGTGGACC-BHQ1) designed by CDC (United States Centers for Disease Control and Prevention). Standard of infectious virus with a known titer was used to quantify the amount of viral RNA (vRNA) as TCID50 (50 % tissue culture infectious dose) per mL. Effect of the compound was expressed as log reduction of vRNA.

Protein expression and purification for crystallisation studies

The 3CLpro of SARS-CoV-2 plasmid was kindly provided by the research group of Prof. Rolf Hilgenfeld from Institute of Biochemistry, Center for Structural and Cell Biology in Medicine, University of Lübeck (Germany) (ORF1ab polyprotein residues 3264-3569, GenBank code: MN908947.3). Protein production and purification was as reported by Zhang et al., 2020 where eluted fractions containing the target protein were pooled and subjected to buffer exchange in 20

mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, 1 mM DTT, pH 7.8.[5] Protein was flash frozen in LN2 at a concentration of 10-20 mg/ml and stored in aliquots at -80°C.

Crystallization, data collection, data reduction, structure determination, refinement and final model analysis.

Prior to crystallization each protein aliquot was thawed in ice and buffer exchanged against 20 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, pH 7.8 to remove DTT excess. Crystals of myricetin: SARS-CoV-2 3CL-Pro were obtained with the vapour diffusion technique, in sitting drops, using seeding. The protein at 5mg/ml was incubated with 30-fold molar excess of myricetin and mixed in equal volume with 0.1M DL-Glutamic acid monohydrate, 0.1M DL-Alanine, 0.1M Glycine, 0.1M DL-Lysine monohydrochloride, 0.1M DL-Serine, 0.1M HEPES/MOPS pH 7.5, 20% v/v Ethylene glycol, 10 % w/v PEG 8000 as a precipitant solution. Diffraction data were collected at 100 K, at the XDR2 beamline of the Elettra Sincrotrone Trieste[45] using a 0.9717 Å wavelength. The collected dataset was processed with XDS [46] and Aimless[47] from the CCP4 suite[48] . The structure was solved by molecular replacement with Phaser[49] using as a search model 7ALH (PDB ID). The initial model was refined alternating cycles of manual model building in COOT[50,51] and automatic refinement using Phenix [52](version 1.18.2_3874).

ACCESSION CODES

Coordinates and structure factors were deposited to the Protein Data Bank with accession code 7B3E.

Data analysis

Data analysis of assay development results was performed using GraphPad Prism 8. In test compound screening assays, data was analysed using commercial software (Activitybase, IDBS). Test compound results were normalised relative to respective controls and control well outliers eliminated according to the three-sigma method. Dose response curves were fitted to 4-parameter logistic functions in the ActivityBase software (ActivityBase Version 8.0.5.4) Assay quality was assessed using the Z'-factor calculation with Z'> 0.5 as threshold for acceptance[53]. TIBCO Spotfire Analyst 7.11 was used for visualization/selection. PDB website ligand interaction scripts (or CCG-MOE version 2019) were used for structural visualizations. By applying a molecular mechanics method integrated with Wavefunction Spartan '18 parallel suite [Version 18.4.1], the equilibrium geometry, among the thioguanosine tautomers has been obtained.

A geometrical docking method was performed with LiGen™, proprietary software developed by Dompé, was used for the identification of the best binding mode[54,55].

Acknowledgements

References

[1] Wu, C., Liu, Y., Yang,Y., Zhang, P., Zhong, W, Wang, Y., Wang, Q., Xu, Y., Li, M., Li, x., Zheng, M., Chen, L., Li, H. (2020). Analysis of therapeutic targets for SARS-CoV-2 and discovery of potential drugs by computational methods. *Acta Pharmaceutica Sinica B, 10, 5, 766-788*. DOI: 10.1016/j.apsb.2020.02.008 ; Hilgenfeld, R. (2014), From SARS to MERS: crystallographic studies on coronaviral proteases enable antiviral drug design. FEBS J, 281: 4085-4096. DOI: 10.1111/febs.12936.

[2] Zumla, A., Chan, J.F.W., Azhar, E.I., Hui, D.S.C., Yuen, K.-Y. (2016). Coronaviruses — drug discovery and therapeutic options. *Nature Reviews Drug Discovery 15(5), 327-347*. DOI:10.1038/nrd.2015.37

[3] WHO COVID19 website https://covid19.who.int/ (latest seen on Nov 18th, 2020).

[4] Cannalire, R., Cerchia, C., Beccari, A.R., Saverio Di Leva, F., Summa, V., (2020). Targeting SARS-CoV-2 Proteases and Polymerase for COVID-19 Treatment: State of the Art and Future Opportunities. Journal of Medicinal Chemistry Article. DOI: 10.1021/acs.jmedchem.0c01140.

[5] Zhang, L., Lin,D., Sun, X., Curth, U., Drosten, C., Sauerhering, L., Becker, S., Rox, K., Hilgenfeld, R., (2020). Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved α-ketoamide inhibitors *Science 368(6489), 409–412.* DOI:10.1126/science.abb3405.

[6] Jin, Z., Du, X., Xu, Y., Deng, Y., Liu, M., Zhao, Y.et al. (2020). Structure of M^{pro} from COVID-19 virus and discovery of its inhibitors *Nature,* 582, 289-293. DOI: 10.1038/s41586-020-2223-y.

[7] Sirois, S., Zhang, R., Gao, W., Gao, H., Li, Y., Zheng, H., Wei, D.-Q. (2007). Discovery of Potent Anti-SARS-CoV MPro Inhibitors, *Current Computer-Aided Drug Design 3, 191*. DOI:10.2174/157340907781695440.

[8] Ramajayam, R., Tan, K.-P., Liang, P.-H. (2011). Recent development of 3C and 3CL protease inhibitors for anti-coronavirus and anti-picornavirus drug discovery. *Biochem Soc Trans 1* 39 (5), 1371–1375. DOI: 10.1042/BST0391371.

[9] Ahn, T.-Y., Ha, D. C., 정 영 식 . (2010). Synthesis and Evaluation of Benzoquinolinone Derivatives as SARS-CoV 3CL Protease Inhibitors. *Bulletin of the Korean Chemical Society* [Internet]. 31(1), 87–91. DOI:10.5012/BKCS.2010.31.01.087.

[10] Kuo, C.-J., Liu, H.-G., Lo, Y.-K., Seong, C.-M., Lee, K.-I., Jung, Y.-S. and Liang, P.-H. (2009). Individual and common inhibitors of coronavirus and picornavirus main proteases, *FEBS Letters*, 583. DOI: 10.1016/j.febslet.2008.12.059.

[11] Lee C.-C., Kuo C.-J. Kuo, Hsu M.-F., Liang P.-H, Fang J.-M., Shie J.-J., Wang A. H.-J. (2007). Structural basis of mercury- and zinc-conjugated complexes as SARS-CoV 3C-like protease inhibitors *FEBS Letters* 581, 28, 5454-5458 DOI: 10.1016/j.febslet.2007.10.048.

[12] Jo,S., Kim,s., Kim, D.Y., Kim, M.-S., Shin,D.H. (2020). Flavonoids with inhibitory activity against SARS-CoV-2 3CLpro, Journal of Enzyme Inhibition and Medicinal Chemistry, 35:1, 1539-1544. DOI: 10.1080/14756366.2020.1801672.

[13] Zhu,W., Xu,M. , Chen, C.Z., Guo, H., Shen, M., Hu, X., Shinn, P., Klumpp-Thomas, C., Michael, S.G., Zheng, W. (2020). Identification of SARS-CoV-2 3CL Protease Inhibitors by a Quantitative High-Throughput Screening ACS Pharmacology & Translational Science 3 (5), 1008-1016. DOI:10.1021/acsptsci.0c00108.

[14] Young,B.E., Xiang Ong,S.W., Kalimuddin, S., (2020). Epidemiologic Features and Clinical Course of Patients Infected With SARS-CoV-2 in Singapore. *JAMA.* 323(15), 1488–1494. DOI:10.1001/jama.2020.3204.

[15] Oscanoa,T.J. , Romero-Ortuno,R., Carvajal,A., Savarino,A. (2020). A pharmacological perspective of chloroquine in SARS-CoV-2 infection: An old drug for the fight against a new coronavirus?, *International Journal of Antimicrobial Agents*, 56(3), 106078. DOI:10.1016/j.ijantimicag.2020.106078.

[16] The RECOVERY Collaborative Group, Dexamethasone in Hospitalized Patients with Covid-19 — Preliminary Report (2020). *New England Journal of Medicine.* DOI: 10.1056/NEJMoa2021436.

[17] Beigel, J. H., Tomashek, K. M., Dodd, L. E., Mehta, A. K., Zingman, B, et al. (2020). Remdesivir for the Treatment of Covid-19 — Final Report *N. Engl. J. Med.* 383, 1813-1826. DOI: 10.1056/NEJMoa2007764

[18] Ma, C., Sacco, M.D., Hurst, B. *et al.* (2020). Boceprevir, GC-376, and calpain inhibitors II, XII inhibit SARS-CoV-2 viral replication by targeting the viral main protease. *Cell Res* 30, 678–692. DOI:10.1038/s41422-020-0356-z.

[19] Data can be retrieved in ChEMBLDB version 28
https://www.ebi.ac.uk/chembl/document_report_card/CHEMBL4495564

[20] Ellinger, B., Bojkova, D., Zaliani, A., Cinatl, J., Claussen, C., Westhaus, S., Reinshagen, J., Kuzikov, M., Wolf, M., Geisslinger, G., Gribbon, P., Ciesek, S. (2020). Identification of inhibitors of SARS-CoV-2 in-vitro cellular toxicity in human (Caco-2) cells using a large scale drug repurposing collection. *ResearchSquare* DOI: 10.21203/rs.3.rs-23951/v1 (in revision at Nature Scientific Data)

[21] Su H, Yao S, Zhao W, et al. (2020). Discovery of baicalin and baicalein as novel, natural product inhibitors of SARS-CoV-2 3CL protease in vitro. bioRxiv; DOI: 10.1101/2020.04.13.038687.

[22] Kneller, D.W., Phillips, G., O'Neill, H.M. *et al.* (2020). Structural plasticity of SARS-CoV-2 3CL M^pro active site cavity revealed by room temperature X-ray crystallography. *Nat Commun* 11, 3202. DOI: 10.1038/s41467-020-16954-7.

[23] Grottesi, A., Bešker, N., Emerson, A., Manelfi, C., Beccari, A.R., Frigerio, F., Lindahl, E., Cerchia, C., Talarico, C. (2020). Computational Studies of SARS-CoV-2 3CLpro: Insights from MD Simulations. *Int J Mol Sci.* 21(15), 5346. DOI: 10.3390/ijms21155346.

[24] Gorbalenya, A.E., Snijder, E.J. (1996). Viral cysteine proteinases. *Perspectives in Drug Discovery and Design* 6, 64–86. DOI: 10.1007/BF02174046.

[25] Günther, S., Reinke, P.Y.A., Oberthuer, D., et al. (2020). Catalytic cleavage of HEAT and subsequent covalent binding of the tetralone moiety by the SARS-CoV-2 main protease *bioRxiv* 2020.05.02.043554, bioRxiv 2020.05.02.043554; DOI: 10.1101/2020.05.02.043554.

[26] Báez-Santos, Y.M., St. John, S.E., Mesecar, A.D. (2015). The SARS-coronavirus papain-like protease: Structure, function and inhibition by designed antiviral compounds *Antiviral Research* 115, 21-38 DOI:10.1016/j.antiviral.2014.12.015.

[27] Chu, H.-F., Chen, C.-C., Moses, D.C., Chen, Y.-H., Lin, C.-H., Tsai, Y.-C., Chou, C.-Y. (2018). Porcine epidemic diarrhea virus papain-like protease 2 can be non-competitively inhibited by 6-thioguanine *Antiviral Research* 158, 199-205. DOI:10.1016/j.antiviral.2018.08.011.

[28] Chuang, S., Cheng, S., Tang, H. *et al.* (2018). 6-Thioguanine is a noncompetitive and slow binding inhibitor of human deubiquitinating protease USP2. *Sci Rep* 8, 3102. DOI:10.1038/s41598-018-21476-w.

[29] Swaim, C.D. , Perng, Y.-C., Zhao, X., Canadeo, L.A. , Harastani, H.H., Darling, T.L., Boon, A.C.M., Lenschow, D.J., Huibregtse, J.M. (2020). 6-Thioguanine blocks SARS-CoV-2 replication by inhibition of PLpro protease activities *bioRxiv* 2020.07.01.183020; DOI:10.1101/2020.07.01.183020

[30] Altun, M., Kramer, H.B., Willems, L.I., McDermott, J.L., et al. Activity-Based Chemical Proteomics Accelerates Inhibitor Development for Deubiquitylating Enzymes (2011). *Chemistry & Biology*, 18 (11, 1401-1412 DOI:10.1016/j.chembiol.2011.08.018.

[31] Tian, X., Isamiddinova, N.S., Peroutka, R.J., et al. (2011). Characterization of selective ubiquitin and ubiquitin-like protease inhibitors using a fluorescence-based multiplex assay format *Assay Drug Dev Technol*. 9(2), 165-173. DOI:10.1089/adt.2010.0317.

[32] Słabicki, M., Kozicka, Z., Petzold, G. *et al.* (2020). The CDK inhibitor CR8 acts as a molecular glue degrader that depletes cyclin K. *Nature* 585, 293–297. DOI:/10.1038/s41586-020-2374-x.

[33] Liu, P., Liu, H., Sun, Q,, Liang, H., Li, C., Deng, X., Liu, Y., Lai, L. (2020). Potent inhibitors of SARS-CoV-2 3C-like protease derived from N-substituted isatin compounds, *European Journal of Medicinal Chemistry* 206, 112702 DOI:10.1016/j.ejmech.2020.112702.

[34] Zastavnyĭ, I.V. (1979). Comparative study of the anti-influenzal effectiveness of bonaphthon, oxoline, florenal and rimantadine. *Mikrobiol Zh.* 41(1), 69-74. https://pubmed.ncbi.nlm.nih.gov/440154/

[35] Chauhan, A., Kalra, S. Identification of potent COVID-19 main protease (MPRO) inhibitors from flavonoids, (2020). *PREPRINT* (Version 1) available at Research Square DOI:10.21203/rs.3.rs-34497/v1

[36] Pandey, K.B., Rizvi, S.I. Plant polyphenols as dietary antioxidants in human health and disease. Oxid Med Cell Longev. (2009). 2(5):270-278. DOI: 10.4161/oxim.2.5.9498.

[37] Baell, J.B., (2016) Journal of Natural Products 79 (3), 616-628 DOI:10.1021/acs.jnatprod.5b00947

[38] Schneider, M., Ackermann, K., Stuart, M., (2012). Severe acute respiratory syndrome coronavirus replication is severely impaired by MG132 due to proteasome-independent inhibition of M-calpain. *J Virol.* 86(18), 10112-10122. DOI:10.1128/JVI.01001-12.

[39] Longhitano, L., Tibullo, D., Giallongo, C., et al. (2020). Proteasome Inhibitors as a Possible Therapy for SARS-CoV-2. Int J Mol Sci. 21(10), 3622. DOI:10.3390/ijms21103622.

[40] Grinev, A.N., Pershin, G.N., Stebaeva, L.F. *et al.* (1984). A new antiviral chemotherapeutant, riodoxol. *Pharm Chem J* 18, 208–211. DOI:10.1007/BF00759009.

[41] Corsello, S., Bittker, J., Liu, Z. *et al.* (2017). The Drug Repurposing Hub: a next-generation drug library and information resource. *Nat Med* 23, 405–408. DOI: 10.1038/nm.4306.

[42] Kim, Y., Liu, H., Galasiti Kankanamalage, A.C., Weerasekara, S., Hua, D.H., et al. (2016). Reversal of the Progression of Fatal Coronavirus Infection in Cats by a Broad-Spectrum Coronavirus Protease Inhibitor. *PLOS Pathogens* 12(3): e1005531. DOI: 10.1371/journal.ppat.1005531.

[43] Joshi,R.S., Jagdale,S.S., Bansode,S.B., Shiva Shankar,S., Tellis,M.B., Pandya, V.K., Chugh,A., Giri,A.P., Kulkarni, M.J. (2020). Discovery of potential multi-target-directed ligands by targeting host-specific SARS-CoV-2 structurally conserved main protease, *Journal of Biomolecular Structure and Dynamics*, DOI: 10.1080/07391102.2020.1760137.

[44] Ivens, T., Van den Eynde, C., Van Acker, K., et al. (2005). Development of a homogeneous screening assay for automated detection of antiviral agents active against severe acute respiratory syndrome-associated coronavirus. *J Virol Methods*. 129(1), 56-63. DOI:10.1016/j.jviromet.2005.05.010.

[45] A. Lausi, M. Polentarutti, S. Onesti, J. R. Plaisier, E. Busetto, G. Bais, L. Barba, A. Cassetta, G. Campi, D. Lamba, A. Pifferi, S.C. Mande, D.D. Sarma, S.M. Sharma, G.Paolucci. (2015). Status of the crystallography beamlines at Elettra *The European Physical Journal Plus*, 130:43,1-8. DOI: 10.1140/epjp/i2015-15043-3.

[46] Kabsch W., (2010). XDS, *Acta Cryst*.D66: 125-132.

[47] Evans P.R, Murshudov G.N., (2013). How Good Are My Data and What Is the Resolution *Acta Crytsallographica* D69: 1204-1214.

[48] Winn M.D., Ballard C.C., Cowtan K.D., Dodson E.J., Emsley P., Evans P.R., Keegan R.M., Krissinel E.B., Leslie A.G., McCoy A., McNicholas S.J., Murshudov G.N., Pannu N.S., Potterton E.A., Powell H.R., Read R.J., Vagin A., Wilson K.S. (2011). Overview of the CCP4 suite and current developments. *Acta Crystallographica* D67: 235-242

[49] McCoy A.J., Grosse-Kunstleve R.W., Adams P., Winn M.D., Storoni L.C. Read R.J., Phaser crystallographic software (2007). *Journal of Applied Crystallography* 40, 658–674.

[50] Emsley P. and Cowtan K. (2004). Coot: model-building tools for molecular graphics." *Acta Crystallographica* D60: 2126-2132.

[51] Emsley P., Lohkamp B., Scott W.G. and Cowtan K. (2010). "Features and Development of Coot." *Acta Crystallographica* D66, 486-501.

[52] Liebschner D., Afonine P.V., Baker M.L., Bunkóczi G., Chen V.B., Croll T.I., Hintze B., Hung L.-W., Jain S., McCoy J., Moriarty N.W., Oeffner R.D., Poon B.K., Prisant M.G., Read R.J., Richardson J.S., Richardson D.C., Sammito M.D., Sobolev O.V., Stockwell D.H., Terwilliger T.C., Urzhumtsev A.G., Videau L.L., Williams C.J., Adams P. (2019). Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallographica* D75, 861–877.

[53] Iversen, P.W., Eastwood, B.J., Sitta Sittampalam, G., Cox, K.L. (2006). A Comparison of Assay Performance Measures in Screening Assays: Signal Window, Z' Factor, and Assay Variability Ratio *Journal of Biomolecular Screening,* 11 (3) 247-252, DOI:10.1177/1087057105285610.

[54] Beccari, A.R., Cavazzoni, C., Beato, C., Costantino, G. (2013). LiGen: a high performance workflow for chemistry driven de novo design. *J Chem Inf Model*. 53(6), 1518-1527. DOI:10.1021/ci400078g.

[55] Beato, C., Beccari, A.R., Cavazzoni, C., Lorenzi, S., Costantino, G. (2013). Use of experimental design to optimize docking performance: the case of LiGenDock, the docking module of LiGen, a new de novo design program. *J Chem Inf Model*. 53(6), 1503-1517. DOI: 10.1021/ci400079k.
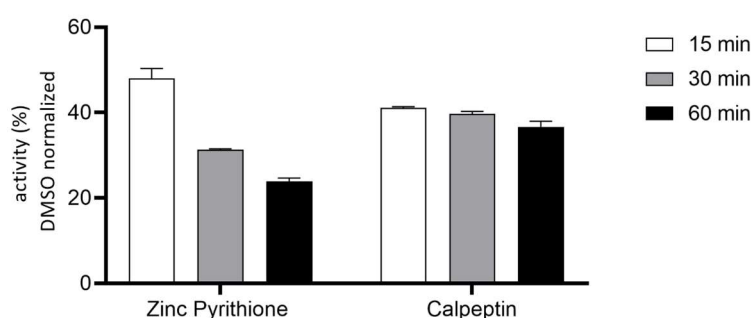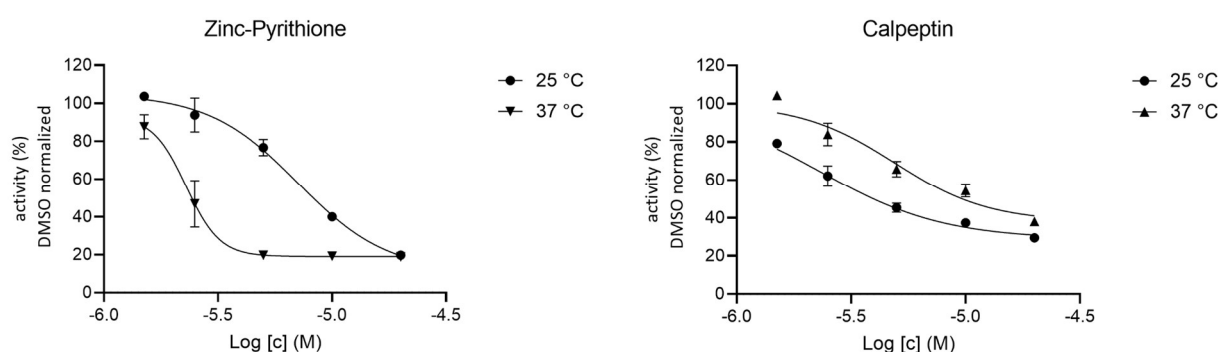
Supplementary Figures
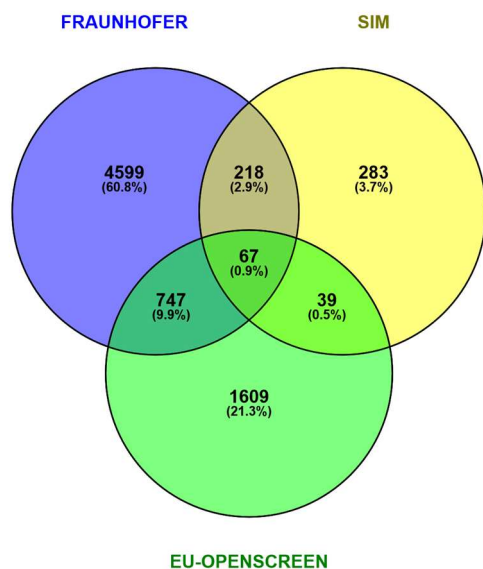
Supplementary Figure 1. Venn diagram showing overlap across the three compound collections (EU-OPENSCREEN, DOMPE Safe in Man (SIM) and EU-OPENSCREEN Bioactives) which composed the screened set. Percentages are relative to the total of 8702 compounds.
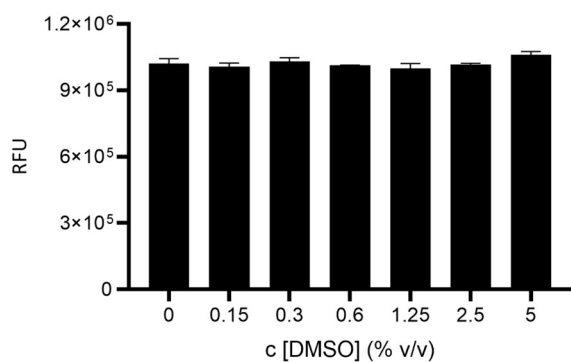

Supplementary Figure 2. Fluorescence intensity versus DMSO concentration % vol/vol. Assay conditions as per SARS-CoV-2 3CL-Pro primary screen.
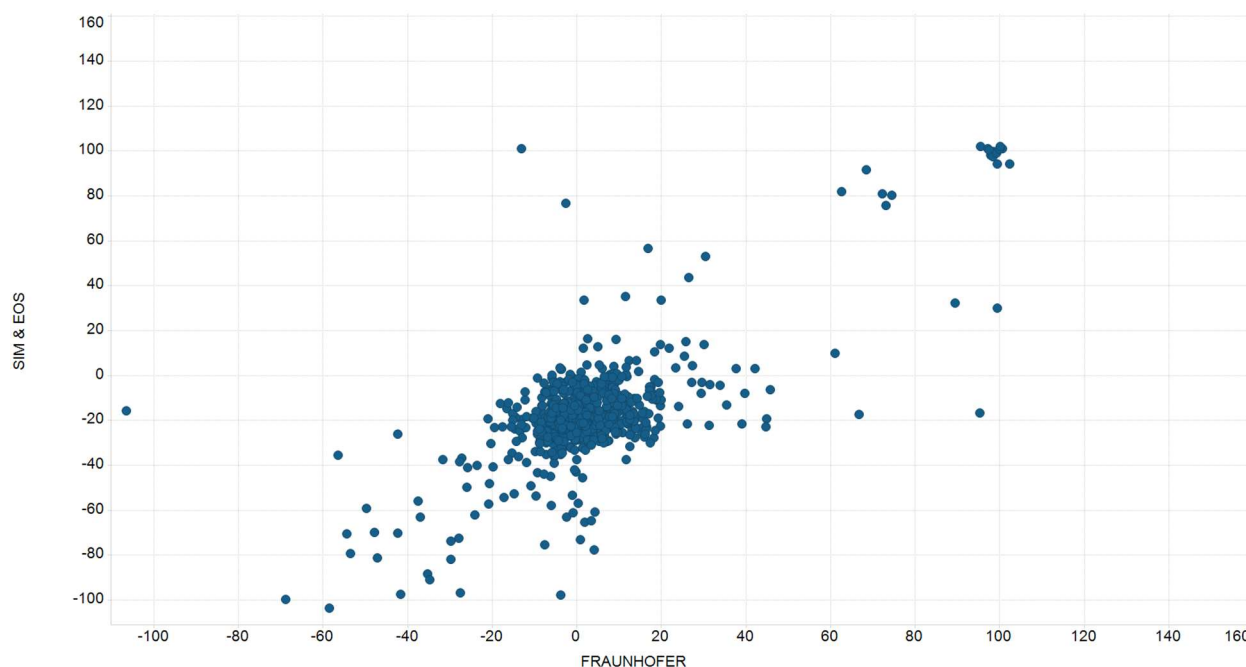
Supplementary Figure 3. Effects of compound pre-incubation times with SARS-CoV-2 3CL-Pro, on inhibitory activity of zinc pyrithione and calpeptin. Data are expressed as measured enzyme activity (%) in the presence of compound, normalised to corresponding DMSO control (100 % activity). Enzyme and substrate concentrations as per primary assay (no DTT). Pre-incubation temperature was 25 °C, with plates read at 15 minutes post substrate addition.


Supplementary Fig 4. Effects of temperature during the pre-incubation stage of reference compounds with SARS-CoV-2 3CL-Pro enzyme. Zinc pyrithione IC_{50}'s = 7.1 and 2.3 µM at 37 °C and 25°C, respectively. Calpeptin IC_{50}'s = 2.3 and 4.9 µM at 37 °C and 25°C, respectively. Assay buffer, enzyme, substrate and timings as per primary assay. SARS-COV-2 3CL-Pro incubated with 0.5 v/v % DMSO represents 100 % activity.
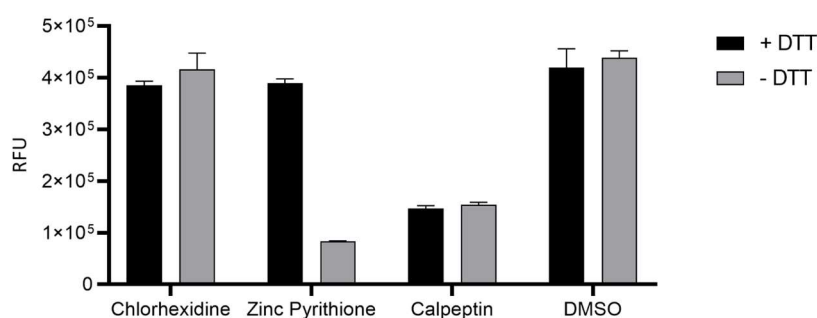
Supplementary Fig 5. Dependence of inhibition of proposed SARS-CoV-2 3CL-Pro inhibitors on DTT (at 0 or 1mM). Test compound concentration 20 µM. Enzyme, substrate and timing conditions as per primary screen.

Supplementary Figure 6. Primary screen - comparison of compounds duplicated in different collections. Primary Screening Inhibition (%) of 3CL-Pro collected by compounds in Fraunhofer repurposing collection (x-axis) versus structurally identical compounds found in EU-OS or DOMPE collections. Data from compounds doubly or triply represented in the libraries (see Figure S1) Linear regression fit R^2= 0.78.
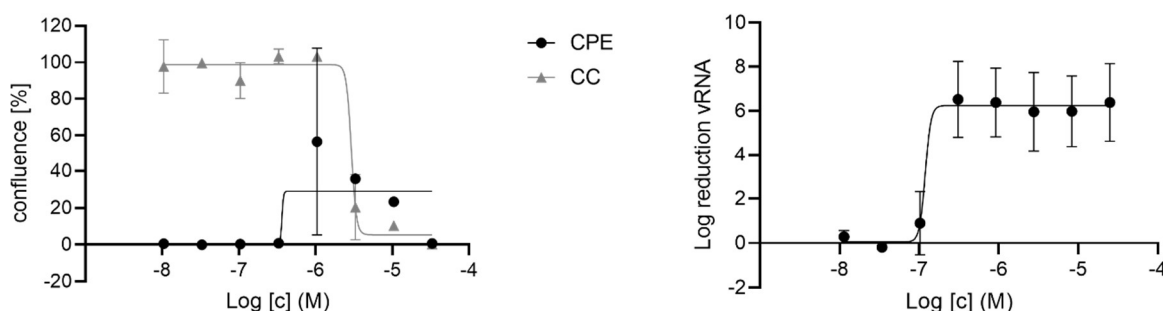
Supplementary Fig 7. MG-132, inhibition of virus induced cytopathic effect in Vero-E6 cells (IC_{50} = 0.36 µM), including the cytotoxicity (CC_{50}= 2.9 µM) versus inhibition of viral replication in Vero-E6 (IC_{50} = 0.12 µM) (assessed by qPCR)
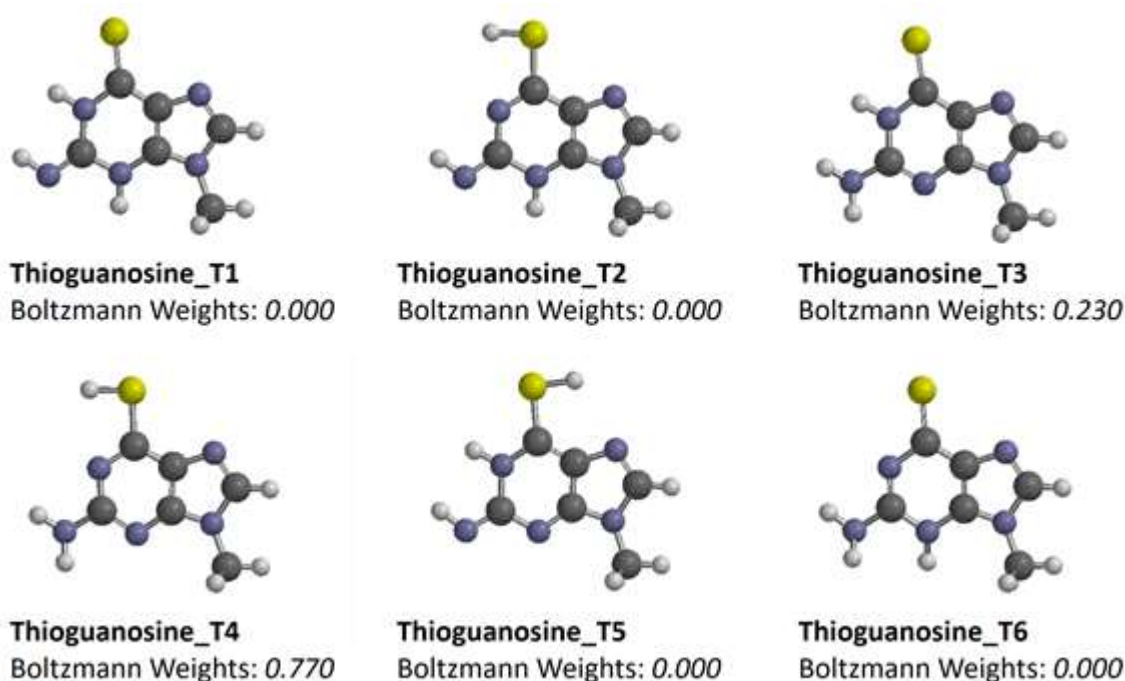
Supplementary Modelling

Due to the thiol group and thio-ketonic group can affect the binding mode of thioguanosine, explaining its reactivity, the tautomeric distributions were calculated, using DFT implemented in Wavefunction Spartan '18, at B3LYP/AUG-cc-pVTZ basis set. Tautomer distribution results showed

a clear separation in terms of Boltzmann weights for the six generated tautomers (Fig S7), highlighting the prevalence of the thioguanosine_T4 form with respect to the thio-ketonic group form thioguanosine_T3.

In particular, the molecular mechanically optimized equilibrium geometry with DFT, using the augmented representations of B3LYP/cc-pVTZ polarization basis sets (AUG-cc-pVTZ) was evaluated [Thom H. Dunning Jr. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *The Journal of Chemical Physics* 90:2, 1007-1023] .

To avoid conformational bias that could occur given the presence of the b-D-ribofuranosyl group, the nitrogen atom at position 6 was methylated.

Thioguanosine_T1
Boltzmann Weights: *0.000*

Thioguanosine_T2
Boltzmann Weights: *0.000*

Thioguanosine_T3
Boltzmann Weights: *0.230*

Thioguanosine_T4
Boltzmann Weights: *0.770*

Thioguanosine_T5
Boltzmann Weights: *0.000*

Thioguanosine_T6
Boltzmann Weights: *0.000*

Supplementary Figure 8. The six optimized structures of thioguanosine tautomers at B3LYP/AUG-cc-pVTZ basis set.

A geometrical docking method was performed with LiGen™, proprietary software developed by Dompé, for the identification of the best binding mode.[54,55] In particular, the docking search was focused within the 3CL-Pro binding site, by defining the free points within a three-dimensional grid, which include the entire binding site. The free points will be used by the docking procedures to define the pharmacophoric key points. The docking software follows a specific workflow during which three docking scores are computed: first, the Pacman Score (PS) estimates a geometric fitting score to evaluate the interaction between a ligand arrangement and the pocket, based on shape and volume complementarity. Then, the Chemical Score (CS), which encodes for the ligand binding interaction energy and is calculated by using an in-house developed scoring function[55]. A rigid body minimization of the docked ligand within the binding site is the last step, at the end of which a third score called the Optimized Chemical Score (Csopt) is evaluated.

Supplementary Tables

Supplementary Table 1. Primary screen quality control. Z prime versus plate ID.

Plate Id	Z Prime
ESP0026074	0.89
ESP0026075	0.86
ESP0026076	0.87
ESP0026077	0.83
ESP0026078	0.91
ESP0026079	0.82
ESP0026080	0.85
ESP0026081	0.84
ESP0026082	0.87
ESP0026083	0.86
ESP0026084	0.87
ESP0026085	0.88
ESP0026086	0.86
ESP0026087	0.88
ESP0026088	0.87
ESP0026089	0.87
ESP0026154	0.69
ESP0026155	0.76
ESP0026478	0.89
ESP0026479	0.91
ESP0026480	0.86
ESP0026481	0.87
ESP0026482	0.91
ESP0026483	0.9
ESP0026484	0.76

Supplementary Table 2. Primary screening results.
(https://www.ebi.ac.uk/chembl/document_report_card/CHEMBL4495564)

Supplementary Table 3. Hit confirmation and profiling data summary

PAPER_reference_S
upplTables_final.xls