



Research Article

Identifying SARS-CoV-2 antiviral compounds by screening for small molecule inhibitors of Nsp3 papain-like protease

Chew Theng Lim^{1*}, Kang Wei Tan^{1*}, Mary Wu², Rachel Ulferts³, Lee A. Armstrong⁴, Eiko Ozono¹, Lucy S. Drury¹, Jennifer C. Milligan¹, Theresa U. Zeisner⁵, Jingkun Zeng¹, Florian Weissmann¹, Berta Canal¹, Ganka Bineva-Todd⁶, Michael Howell², Nicola O'Reilly⁶, Rupert Beale³, Yogesh Kulathu⁴,  Karim Labib⁴ and  John F.X. Diffley¹

¹Chromosome Replication Laboratory, The Francis Crick Institute, 1 Midland Road, London NW1 1AT, U.K.; ²High-Throughput Screening, The Francis Crick Institute, 1 Midland Road, London NW1 1AT, U.K.; ³Cell Biology of Infection Laboratory, The Francis Crick Institute, 1 Midland Road, London NW1 1AT, U.K.; ⁴MRC Protein Phosphorylation & Ubiquitylation Unit, School of Life Sciences, University of Dundee, Dow Street, Dundee DD1 5EH, U.K.; ⁵Cell Cycle Laboratory, The Francis Crick Institute, 1 Midland Road, London NW1 1AT, U.K.; ⁶Peptide Chemistry STP, The Francis Crick Institute, 1 Midland Road, London NW1 1AT, U.K.

Correspondence: John F.X. Diffley (john.diffley@crick.ac.uk)



The COVID-19 pandemic has emerged as the biggest life-threatening disease of this century. Whilst vaccination should provide a long-term solution, this is pitted against the constant threat of mutations in the virus rendering the current vaccines less effective. Consequently, small molecule antiviral agents would be extremely useful to complement the vaccination program. The causative agent of COVID-19 is a novel coronavirus, SARS-CoV-2, which encodes at least nine enzymatic activities that all have drug targeting potential. The papain-like protease (PLpro) contained in the nsp3 protein generates viral non-structural proteins from a polyprotein precursor, and cleaves ubiquitin and ISG protein conjugates. Here we describe the expression and purification of PLpro. We developed a protease assay that was used to screen a custom compound library from which we identified dihydrotanshinone I and Ro 08-2750 as compounds that inhibit PLpro in protease and isopeptidase assays and also inhibit viral replication in cell culture-based assays.

Introduction

By early January 2021, COVID-19 infections and deaths across 191 countries/regions were reported to be 87 million and 1.9 million, respectively [1]. These numbers occurred in just over one year since the first COVID-19 outbreak was reported in December 2019. COVID-19 is a highly contagious disease with a human-to-human mode of transmission [2]. Accelerated by travelers, by the end of January 2020, the World Health Organization (WHO) declared the disease as a 'public health emergency of international concern' and the first European COVID-19 case was reported in France [3].

The causative agent of COVID-19 was identified as a novel coronavirus, Severe Acute Respiratory Syndrome coronavirus 2 (SARS-CoV-2). The SARS-CoV-2 genome was first sequenced and published in early January 2020 [4]. COVID-19 falls within the subgenus *Sarvecovirus*, genus *Betacoronavirus*, and is closely related (88% nucleotide sequence identity) to a bat-derived SARS-like coronavirus, and to the virus causing SARS (SARS-CoV-1; 79%), and more distantly related to MERS-CoV (50%) the causative agent of MERS (Middle East Respiratory Syndrome) [5].

There is currently no cure for COVID-19. In June 2020, dexamethasone, an anti-inflammatory drug, was approved by the UK government for COVID-19 patients in National Health Service (NHS) hospitals [6]. In August 2020, remdesivir was approved by the Food and Drug Administration (FDA)

*These authors contributed equally to this work.

Received: 6 April 2021
Revised: 5 May 2021
Accepted: 7 June 2021

Accepted Manuscript online:
8 June 2021
Version of Record published:
2 July 2021

in the United States [7]; this is currently the only small molecule antiviral drug approved for use against COVID-19. In addition, casirivimab and imdevimab, monoclonal antibody drugs, were approved under Emergency Use Authorization (EUA) by the FDA in November 2020 [8]. The other focus of treatment is prevention via vaccination. Multiple mRNA-based and virus-based vaccines have been rolled out across the world with similar overall safety and effectiveness [9–11]. However, the emergence of a new variant of SARS-CoV-2 in the UK in September 2020, has raised great concern since it is 70% more transmissible [12]. Additional variants arising across the world may render the existing vaccines less efficient. Given the uncertainty of the progression of the virus and the time frame needed to vaccinate the global community, it is crucial to search for drugs to provide treatment for COVID-19 patients.

The genome of COVID-19 is arranged as shown in Figure 1A [4]. Once the virus has entered a cell, two open reading frames, ORF 1a and ORF 1b are translated. ORF 1a is generated from an internal ribosomal frame shift. This produces polypeptides that are processed by two viral proteases to produce 16 non-structural proteins (nsp) 1 to 16. The main protease is the 3C-like protease (3CLpro), corresponding to nsp5. 3CLpro cleaves in between each of the nsp 4–16 proteins. The second protease, Papain-like protease (PLpro), is contained in a small domain of the nsp3 protein; it cleaves after LXGG motifs between each of the remaining

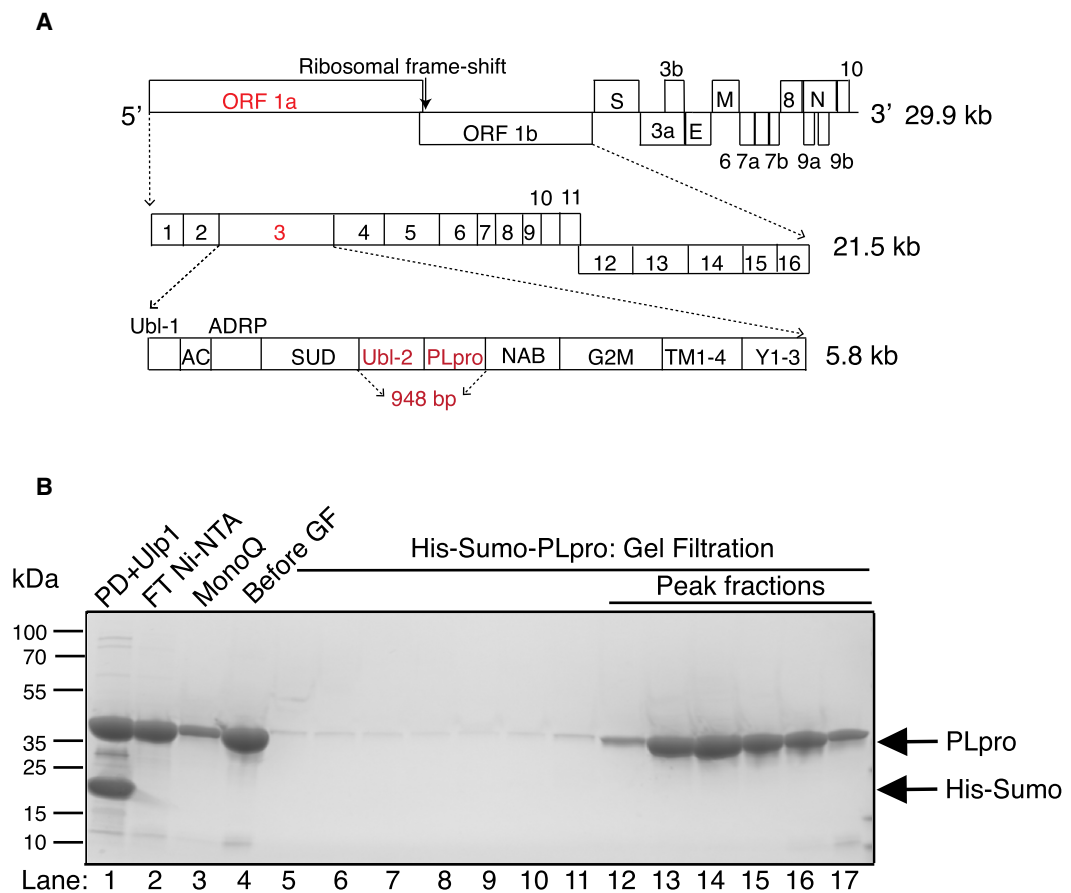


Figure 1. Design, expression and purification of PLpro enzyme.

(A) Schematic of the COVID-19 genome showing the 29.9 kb single-strand (+) RNA. ORF 1a is 21.5 kb and codes for a polyprotein which after processing, produces 16 proteins named nsp1–16. The nsp3 protein contains the catalytic core PLpro enzyme used in this study which comprises the Ubl2 and PLpro domains (highlighted in red). The ‘948 bp’ highlighted in red indicates the sequence that was the basis of the bacterial and insect cell expression constructs used for protein purification.

(B) Purification of His-Sumo-PLpro. Lane 1, Digestion of the pull-down (PD) with Ulp1; lane 2, flow-through (FT) of second Ni-NTA; lane 3, FT from the MonoQ; lane 4, concentrated from lane 3 before applying to a gel filtration column; lanes 13–16 are the gel filtration fractions containing PLpro monomer from that were pooled for further assay.

nsp1–3 [13]. Interestingly, PLpro is also a deubiquitylase (DUB) and a delSGylase, cleaving after the diglycine residues of ubiquitin (Ub) and the UBL protein ISG15 (interferon-induced gene 15), respectively, and thus has roles outside polyprotein cleavage [14,15]. In this paper, we describe our search for PLpro inhibitors.

Results

PLpro expression and purification

Nsp3 is the largest non-structural protein (1945 amino acids). It contains multiple domains which are arranged in the following order: ubiquitin-like (Ubl-1), acidic-domain (AC domain), ADP-ribose-1'-phosphatase (ADRP)/macro/x-domain, SARS unique Domain (SUD), Ubl-2, PLpro domain, nucleic acid-binding domain (NAB), marker domain (G2M), double-pass transmembrane domains (TM 1–2 and TM 3–4), and the Y domain (subdomains Y1–3) [14,16] (Figure 1A). For expression and purification, we selected the region from 1564–1878 amino acids that has been previously described to produce a truncated nsp3 which encompasses the Ubl-2 and PLpro main domains [4,14].

We expressed PLpro in both bacterial and insect cell systems to determine which versions of the purified protein retained the most enzymatic activity. For bacterial expression, PLpro was tagged with either His-Sumo or His-TEV at its N-terminus. After protein pulldown from bacterial lysate, the His-Sumo and His-TEV affinity tags were removed by Ulp1 and TEV proteases respectively (Figure 1B and Supplementary Figure S1A). For expression in insect cells, the protein was tagged with Flag-His at the N-terminus and the tag remained on the final protein (Supplementary Figure S1B). The activities of all three proteins were compared as described below.

PLpro protease activity

We used a quenched Förster (fluorescence) resonance energy transfer (FRET) technique to monitor the protease activity [17]. 2-aminobenzoyl (Abz) and a nitro-L-tyrosine (Y-(3-NO₂)R), were added to opposite ends of a small synthetic peptide that contained the cleavage sequence recognized by PLpro. This generates a fluorescence-quenching pair (Figure 2A) in which emission from Abz is absorbed by the neighboring Y-(3-NO₂)R; cleavage is detected as an increase in apparent emission from Abz at 420 nm. Two cleavage sequences were initially selected which corresponded to the ten amino acids between nsp1/2 and nsp2/3, designated Pro1 and Pro2, respectively. To test for cleavage of the substrates, bacterial His-TEV-PLpro was added to the substrate in the assay buffer. Over the 20 min of incubation, increasing fluorescence signal was observed for substrate Pro2 but not for Pro1 (Supplementary Figure S2A). Consistent with this, it has been recently shown that full-length nsp3 protein but not the isolated PLpro domain can cleave between nsp1/2 [18]. Although cleavage of Pro2 occurred, the rate was relatively slow making it less suitable for screening. We therefore designed a third substrate based on Pro2, called Pro3, which lengthened the recognition sequence peptide from 10 to 12 amino acids [19] (Figure 2A). This modification resulted in a more rapid reaction (Supplementary Figure S2B) and Pro3 was, therefore, used as substrate in subsequent experiments.

The bacterial His-TEV-PLpro (tag removed) was more active than insect cell Flag-His-PLpro at the same enzyme concentration (Supplementary Figure S2D). The two bacterial expression proteins, His-TEV and His-Sumo-PLpro (tags removed) showed similar activity (Supplementary Figure S2C). Since the His-Sumo PLpro is an intact protein with no extra amino acids (after tag removal) and also produced a higher yield after purification, this version of PLpro was used for all of the remaining experiments.

By assaying PLpro cleavage activity across a wide range of substrate concentrations, we found that reaction velocities did not approach saturation at the highest concentration tested and therefore the apparent K_M was estimated to be 1854 μ M based on an incomplete dataset (Figure 2B). Cleavage increased over an hour at a variety of enzyme concentrations (Figure 2C). We chose 1.75 μ M of enzyme and 20 μ M of the substrate to use in the high-throughput screen.

The high-throughput screening for over 5000 compounds

The screen to identify small molecule inhibitors was performed using an existing custom library of over 5000 compounds aliquoted at two different concentrations, 1.25 μ M or 3.75 μ M, in a total of 48 384-well plates (see [20] this issue for contents and description of the library). The plates were organized with columns 3–22 containing the compounds; the remaining columns were used for reaction controls including substrate only and substrate with enzyme but no compound (Figure 3Aii). The PLpro enzyme was pre-incubated with the

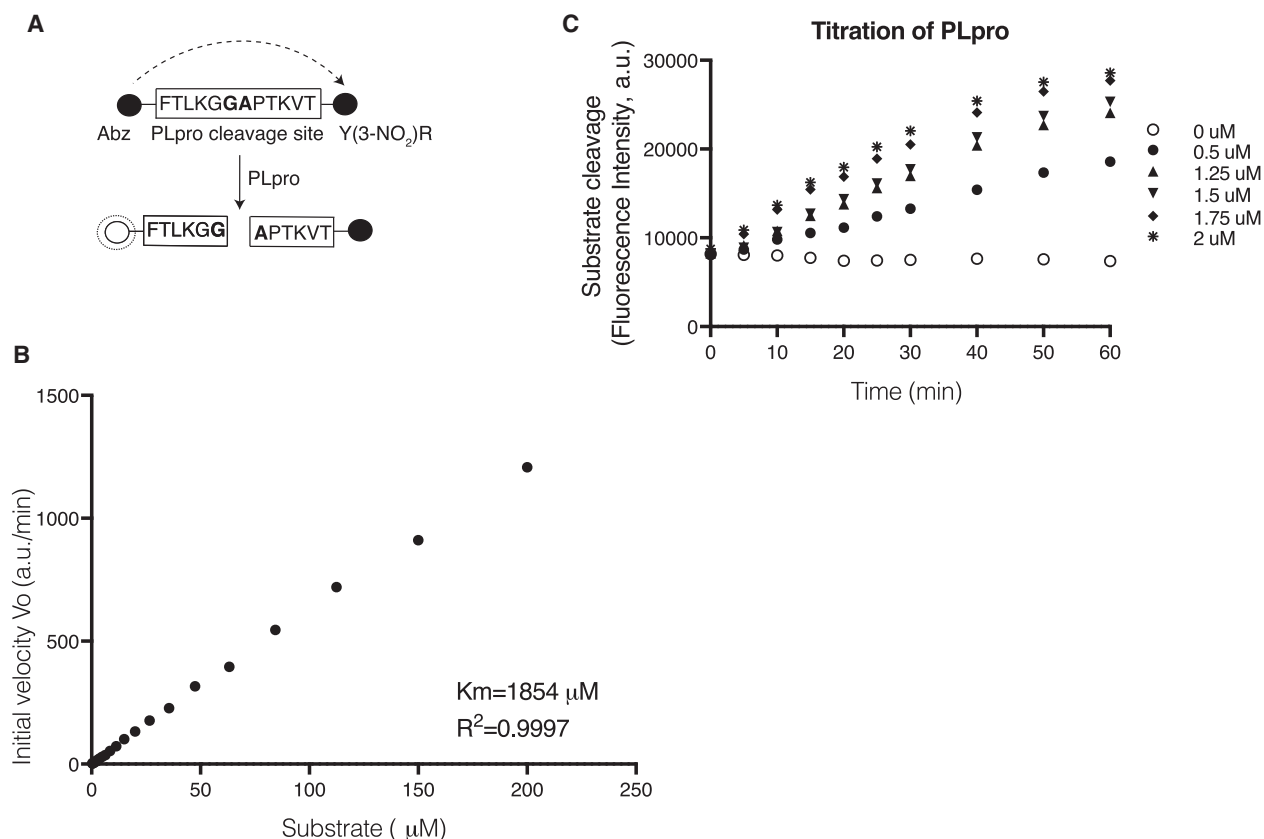


Figure 2. Enzyme assay design and enzyme characteristic.

(A) Schematic of the Pro3 peptide designed for used in the FRET assay. The synthetic peptide contains 12 amino acids from the nsp2/3 junction in the natural polypeptide, in the centre of this is the FTLKGG/APTKVT sequence recognized by PLpro which cleaves between G and A. During synthesis the peptide had the fluorescent Anthranilate (2-aminobenzoyl-Abz) tag added to the N-terminus and the quencher nitro-L-tyrosine (Y (3-NO₂)R) fused to the C-terminus. (B) Determination of the enzyme kinetics. The initial velocity of substrate hydrolysis over the titration of substrate is plotted. The velocity did not saturate at the tested concentration, so K_M were estimated using Michaelis–Menten equation based on the incomplete dataset. Data was collected from three replicates. (C) Protease activity of PLpro. Titration of the purified enzyme (0.5–2 μM) incubated with the substrate. Fluorescent intensity was measured for 1 h. Data was collected from three replicates.

compounds for 10 min, and then the reaction was initiated by the addition of substrate. The activity was recorded from 0 to 20 min with 3 min intervals between readings, this resulted in seven timepoints which were then used to calculate the slope (Figure 3Ai). We incubated the inhibitors with PLpro before the enzymatic reaction so that the slow-binding inhibitors will not be missed out [21].

We selected a total of 29 candidates from the plates using 3.75 μM compound concentration and a further 8 candidates were selected from the plates with 1.25 μM concentration (Figure 3B) that exhibited apparent inhibition greater than 25%; four of the compounds were overlapping in both concentrations. Amongst the hits, 22 out of 29 compounds from the high concentration plates and 1 compound from the low concentration plates were auto-fluorescent at 420 nm, and were excluded from our list. Three additional candidates were eliminated because the degree of inhibition was weak and the same compounds did not inhibit at higher concentration. This narrowed the final list down to seven compounds (Table 1). The details of normalized activity to control without compound of each compounds were stated in Supplementary Table S1.

Gel-based PLpro assay to test candidates

We wanted to exclude artifacts that might result from the fluorescent-based assay. For this reason we designed a gel-based PLpro protease assay (Figure 4A,B). We constructed a new version of the PLpro substrate that had the 10 amino acids at the junction of nsp2–3 containing the cleavage site attached to GST at the N-terminus

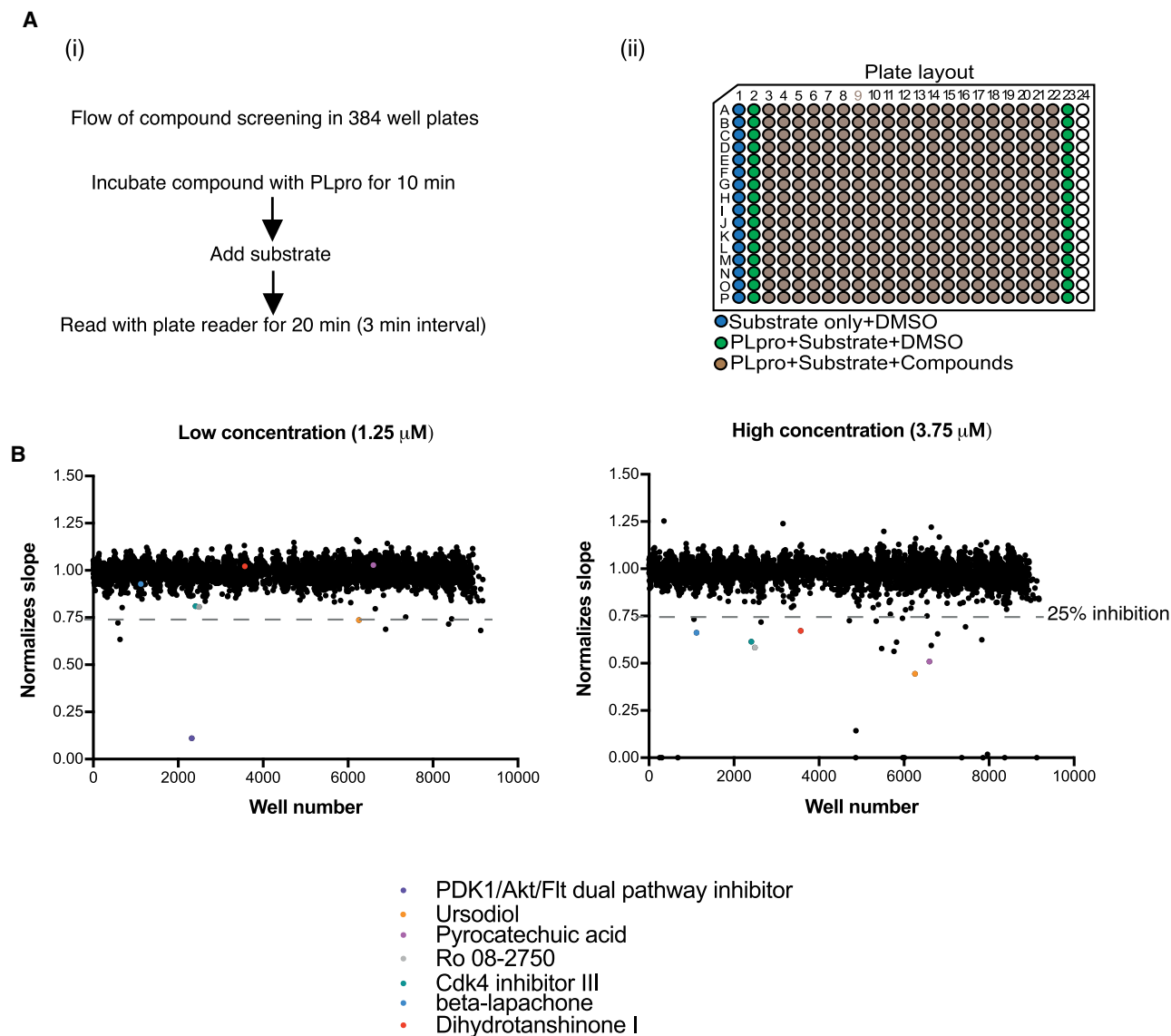


Figure 3. High-throughput screening of the compound library.

(A) (i) Flow diagram of the compound screen. Over 5000 compounds were dispensed into 24 of 384-well plates, enzyme was pre-incubated with compound before the addition of substrate and fluorescent intensity was read at an excitation 330 nm and emission at 420 nm with a Tecan Spark plate reader. (ii) Organization in 384-well plate. (B) Results of the screen. Two concentrations of the compound library were aliquoted- 1.25 μ M and 3.75 μ M. A single dot from the scatter plot represents each of the over 5000 compounds. There are 8 compounds and 29 compounds, in low and high concentration, respectively that reduce the PLpro activity more than 25%. Amongst those hits, four compounds were overlapping. In the low concentration (1.25 μ M), one compound is auto-fluorescent and three have a Z score of more than -3.5 .

and MBP at its C-terminus, resulting in a 67 kDa peptide (Supplementary Figure S3A). If PLpro mediated cleavage occurs, products of 25 kDa and 42 kDa will be generated which can easily be detected and visualized by SDS-PAGE (see Figure 4Ci, lane 3).

One of the hits from the screen, dihydrotanshinone I was previously shown to inhibit the PLpro from SARS-CoV-1 [21]. Two other tanshinone derivatives, tanshinone IIA and cryptotanshinone, were shown to be better inhibitors of this enzyme than dihydrotanshinone I [21]. In separate studies, a non-covalent inhibitor, GRL-0617, has been reported to be effective against SARS-CoV-1/2 PLpro by different groups [22–24]. We, therefore, decided to include these compounds in our validation experiments. As shown in Figure 4C, five of seven candidate hits — PDK1/Akt/Fit dual pathway inhibitor (Figure 4Cii), Cdk4 inhibitor III (Figure 4Ci),

Table 1 Percentage of inhibition and Z score on hits from screening

No.	Name	Inhibition (%)		Z score
		3.75 μ M (slope value/control)	1.25 μ M (slope value/control)	
1	PDK1/Akt/Flt Dual Pathway Inhibitor	100 (0/491)	89 (50/452)	−14.53
2	Ursodiol	55 (187/420)	26 (324/440)	−7.78
3	Pyrocatechuic acid	49 (226/444)	0 (453/441)	−6.846
4	Ro 08-2750	42 (287/491)	19 (366/452)	−5.79
5	Cdk4 Inhibitor III	39 (302/491)	19 (367/452)	−5.34
6	beta-Lapachone	34 (322/486)	7 (420/453)	−4.66
7	Dihydrotanshinone I	33 (329/490)	0 (449/440)	−4.52

dihydrotanshinone I, Ro 08-2750 (Figure 4Ciii) and beta-lapachone — strongly inhibited the enzyme activity at both compound concentrations tested as indicated by the reduced levels of the 25 and 42 kDa products (red dot in Figure 4). Two of the final seven compounds, ursodiol and pyrocatechuic acid, did not inhibit PLpro cleavage of the substrate in this assay and, therefore, appear to be false positives. We tested some of the auto-fluorescent hits, but none inhibited cleavage of the peptide, further confirming them as false positives (Figure 4Cii,iii). Tanshinone IIA and cryptotanshinone also inhibited cleavage, but only at the higher concentration. Thus, dihydrotanshinone I appears to be the strongest inhibitor amongst the tanshinone derivatives. Similarly, GRL-0617 also inhibited cleavage only at the highest concentration. 3CLpro, contained in the nsp5 protein, is the other main protease encoded by SARS-CoV-2, and, like PLpro, 3CLpro has also an active site cysteine. However, none of the five screen hits inhibited 3CLpro activity in an analogous gel-based assay (Milligan et al. this series; Supplementary Figure S3B).

We determined the IC_{50} value for all 5 inhibitors from the screen that were validated using the gel-based assay. All of them are below 1 μ M: 0.26 μ M for PDK/Akt/Flt dual pathway inhibitor, 0.53 μ M for Ro 08-2750, 0.39 μ M for Cdk4 inhibitor III, 0.61 μ M for beta-lapachone, and 0.59 μ M for dihydrotanshinone I (Figure 5A–E). The IC_{50} for the 3 published compounds were slightly higher in our experimental conditions: 1.79 μ M for GRL-0617, 1.57 μ M for tanshinone IIA, and 1.34 μ M for cryptotanshinone (Figure 5F–H), consistent with the gel-based assay results. During this time we realized that PDK1/Akt/Flt dual pathway inhibitor was also identified in several other screens being performed in the laboratory. We, therefore, considered it to likely to be non-specific and we eliminated it from further consideration.

Inhibition of PLpro isopeptidase activity

We performed orthogonal assays using two different substrates of PLpro to test the inhibition of isopeptidase activity of PLpro by the small molecule inhibitors identified in the screen. When K48-linked triubiquitin (Ub3) is incubated with PLpro, it is efficiently cleaved to diubiquitin (Ub2) and monoubiquitin (Ub1) as the final products [18]. Similarly, pro-ISG15 is cleaved to ISG15. When PLpro is pre-incubated with 10 μ M of beta-lapachone or dihydrotanshinone I, we observed potent inhibition of both K48-linked Ub3 and pro-ISG15 cleavage whereas moderate inhibition was observed for both Ro 08-2750 and the previously identified inhibitor of PLpro, GRL-0617, at these concentrations (Figure 6). A potential 3CLpro inhibitor, shikonin, did not inhibit isopeptidase activity with either substrate. These experiments demonstrate the inhibitory effect of these small molecules across a spectrum of PLpro substrates.

Cell culture-based antiviral proliferation assay

We next tested the ability of these compounds to inhibit viral growth in a cell culture-based assay where VERO E6 cells are infected with SARS-CoV-2. Two of the compounds, beta-lapachone and Cdk4 inhibitor III, were cytotoxic in the low micromolar range and were not pursued (Supplementary Figure S4A,B). The two tanshinone derivatives not identified in our screen (cryptotanshinone and tanshinone IIA) did not inhibit viral growth below 200 μ M (Supplementary Figure S4C,D). Ro 08-2750 and GRL-0617 were better at inhibiting viral growth (EC_{50} 20 and 32.6 μ M, respectively) but also exhibited some cell toxicity at higher concentrations

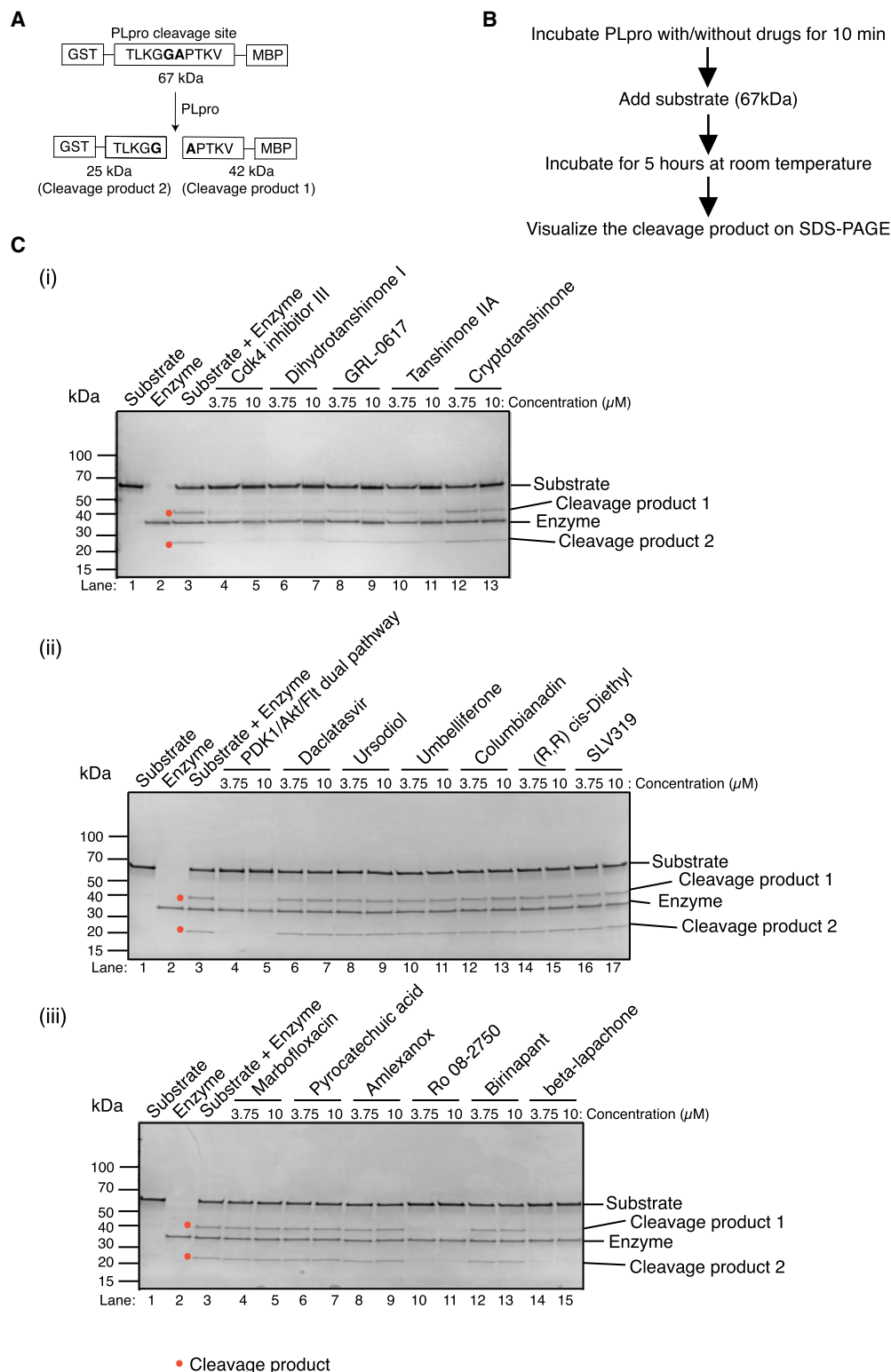


Figure 4. Validation of the hits with a gel-based protease assay.

Part 1 of 2

(A) Schematic of the substrate designed for the gel-based protease assay. A polypeptide with a N-terminal GST and a C-terminal MBP domain separated by 10 amino acids from the natural nsp2/3 junction (similar to that used in the fluorescence assay) was constructed. Cleavage between the amino acids G and A results in 25 kDa and 42 kDa products. (B) Flow diagram of the gel-based assay. (C) (i–iii) Two concentrations of each compound were used, 3.75 μ M and 10 μ M. Only Cdk4 inhibitor III,

Figure 4. Validation of the hits with a gel-based protease assay.

Part 2 of 2

dihydrotanshinone I, GRL-0617, tanshinone IIA, cryptotanshinone, PDK1/Akt/Fit dual pathway inhibitor, Ro 08-2750 (2,3,4,10-Tetrahydro-7,10-dimethyl-2,4-dioxbenzo[g]pteridine-8-carboxaldehyde) and beta-lapachone show inhibition towards PLpro cleavage. GRL-0617, tanshinone IIA, and cryptotanshinone are the published compound inhibitors of PLpro.

(Figure 7Bi,Ci and Supplementary Figure S5). Dihydrotanshinone I proved to be the best inhibitor since it effectively inhibited the SARS-CoV-2 proliferation at an EC₅₀ of 8 μM (Figure 7Ai) and did not exhibit much cytotoxicity, even at high concentrations.

Since remdesivir is the one and only approved antiviral compound for COVID-19, we wanted to test whether any of the compounds we have found show any synergy with remdesivir (Supplementary Figure S5). However, addition of 0.5 μM remdesivir, a concentration just below that required to inhibit viral growth, did not reduce the EC₅₀ of any of our compounds (Figure S7A–Cii). The inhibitory properties of remdesivir alone was described in nsp14 methyltransferase paper in this issue [25].

Discussion

Dihydrotanshinone I, which emerged as the best overall hit from our screen, is a natural compound isolated from lipophilic fraction of *Salvia miltiorrhiza*, which has a long history in traditional Chinese medicine [26]. Several derivatives of tanshinones were previously reported to be inhibitors for SARS-CoV-1 PLpro and to a lesser extent in 3CLpro [21,26]. We have found that dihydrotanshinone I is the best inhibitor of SARS-CoV-2 PLpro and did not inhibit 3CLpro. Although we, like other groups [24,27,28], used a truncated nsp3 protein including just the Ubl2 and PLpro domains, and there is evidence that cleavage by the full-length nsp3 may have slightly different specificities [18]. The fact that dihydrotanshinone I stops viral replication suggests that it is a good nsp3 PLpro inhibitor in cells. Recently, an *in silicon* molecular docking study suggests tanshinone I, derivative of dihydrotanshinone I, directly form the hydrogen bond with the side chain of catalytic C111 amino acid in PLpro [29].

PLpro also recognizes and removes K48-linked polyubiquitin chains (Ub) and ISG15 from host cell target proteins. It is known that either ubiquitin or ISG15 are covalently bonded to target proteins during the cellular response to viral infection. The deubiquitinating (DUB) and deISGylating activities of PLpro following LXGG motifs thus have an implication in viral invasion by shutting down the viral-induced host innate immune response [30]. PLpro of SARS-CoV-2 further shows preferential activity in cleaving ISG15 over Ub in comparison with SARS-CoV-1 [18,22,24]. Interestingly, although the high transmission (and potentially more deadly) U.K. variant of SARS-CoV-2, B.1.1.7 was believed attributed to multiple mutations in the spike protein, the significance of a particular point mutation, A1708D, in PLpro for viral invasion remains unexplored [31].

Experimental procedures

Expression constructs

The coding sequence of SARS-CoV-2 nsp3 1564–1878 amino acid (NCBI reference sequence NC_045512.2) was selected as previously reported [14]. The His-TEV bacterial sequence was codon optimized (Supplementary Table S2) and cloned into plasmid pET11a at NdeI/BamHI sites (synthesized and cloned by GeneWiz) (Addgene ID 169192).

The His-Sumo bacterial version was codon optimized and cloned into K27-Sumo (Addgene ID 169193) via NEBuilder HiFi DNA Assembly Cloning Kit (NEB). Vector template was amplified using primers oEcoli-C_48 and 49, PLpro sequence was amplified from His-TEV bacteria strain using primers oEcoli-C_51 and 52.

Expression and purification

Bacteria

His-TEV and His-Sumo constructs were introduced into T7 express lysY/I^q *E. coli* cell (NEB) for expression. Cells were grown at 37°C to log phase to achieve OD 0.8. Cells were then induced by the addition of 0.5 mM IPTG and switched to 18°C to incubate overnight. Cells were harvested and lysed in buffer A (50 mM Tris-HCl, pH 7.5, 10% glycerol, 1 mM DTT, 0.02% NP-40, 500 mM NaCl and 30 mM imidazole), with the addition of 100 μg/ml lysozyme and sonicated 24 × 5 s. Lysates were centrifuged and the supernatant was collected. The supernatant was incubated with Ni-NTA agarose beads (Thermo) for 2 h at 4°C. Beads were washed with wash

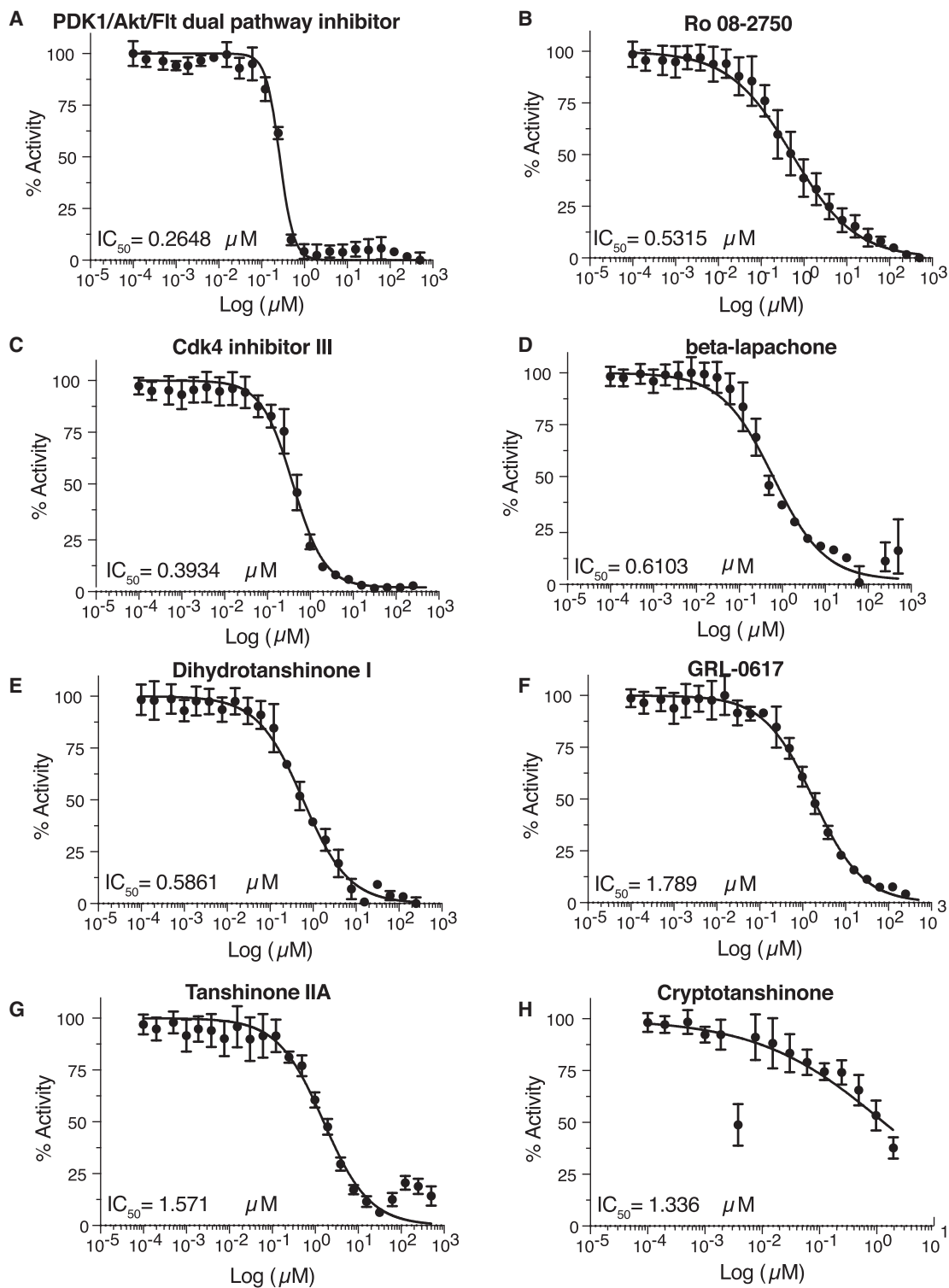


Figure 5. Determination of the IC_{50} of the validated hits with the fluorescent-based assay.

(A) PDK1/AKT/Fit dual pathway inhibitor, (B) Ro 08-2750, (C) Cdk4 inhibitor III, (D) beta-lapachone, (E) dihydratanshinone I, (F) GRL-0617, (G) tanshinone IIA, (H) cryptotanshinone.

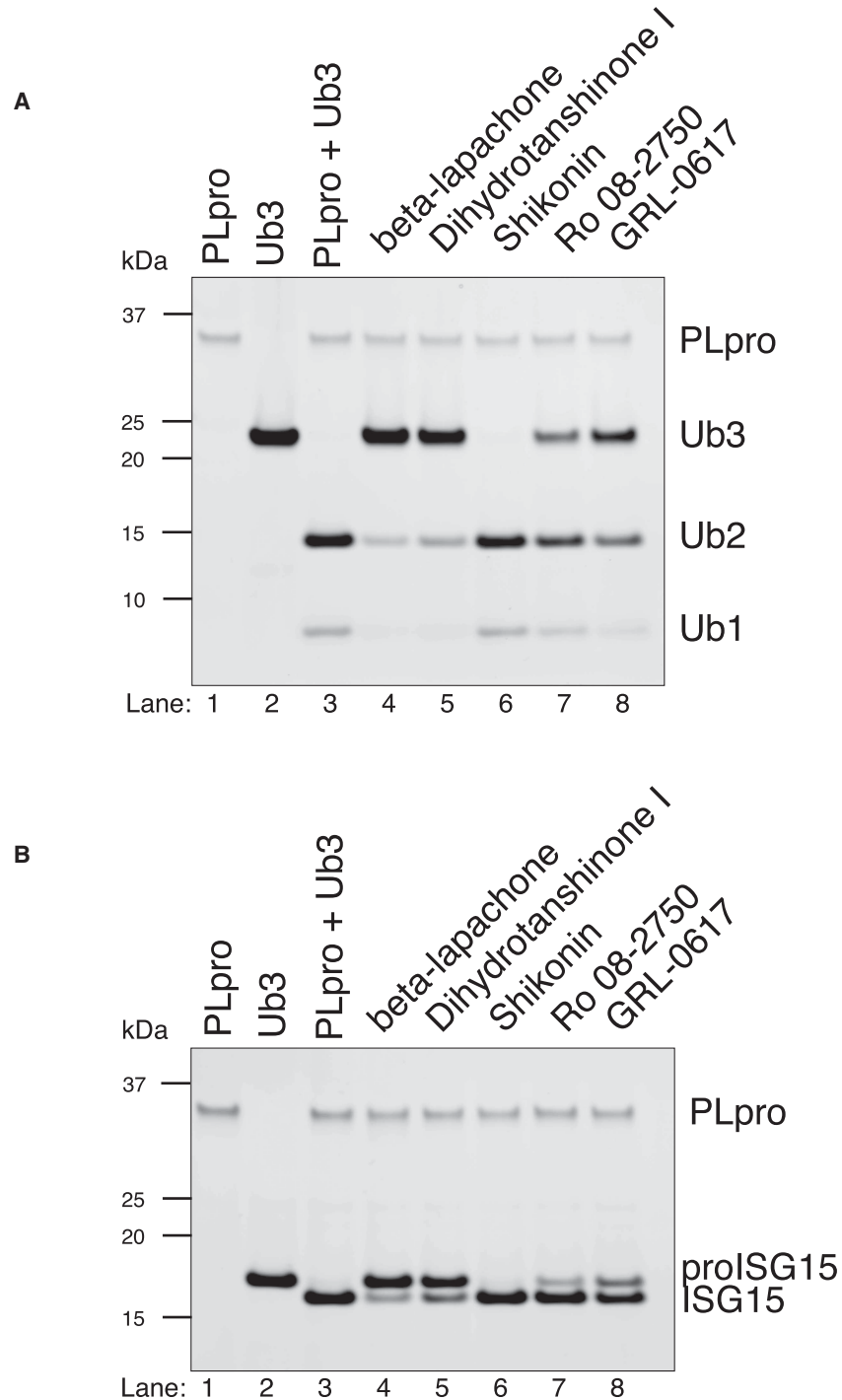


Figure 6. Inhibition of PLpro isopeptidase activities.

An amount of 10 μ M of each compounds were used. **(A)** PLpro DUB activities. PLpro cleaves K48-linked Ub3 to Ub2 and Ub1. These activities were inhibited strongly by beta-lapachone and dihydrotanshinone I, moderately by Ro 08-2750 and GRL-0617, but not by shikonin. **(B)** PLpro delSGylating activities. Pro-ISG15 is being cleaved to ISG15 by PLpro. Similar observation as in **A**.

buffer A. The protein was eluted with 200 mM (His-TEV) or 400 mM (His-Sumo) of imidazole. Fractions were pooled and dialyzed in buffer B (50 mM Tris-HCl, pH 7.5, 10% glycerol, 1 mM DTT, 0.02% NP-40 and 50 mM NaCl) and 0.1 mg/ml His-TEV protease (His-TEV) or 0.02 mg/ml His-Ulp1 (His-Sumo) to cleave-off

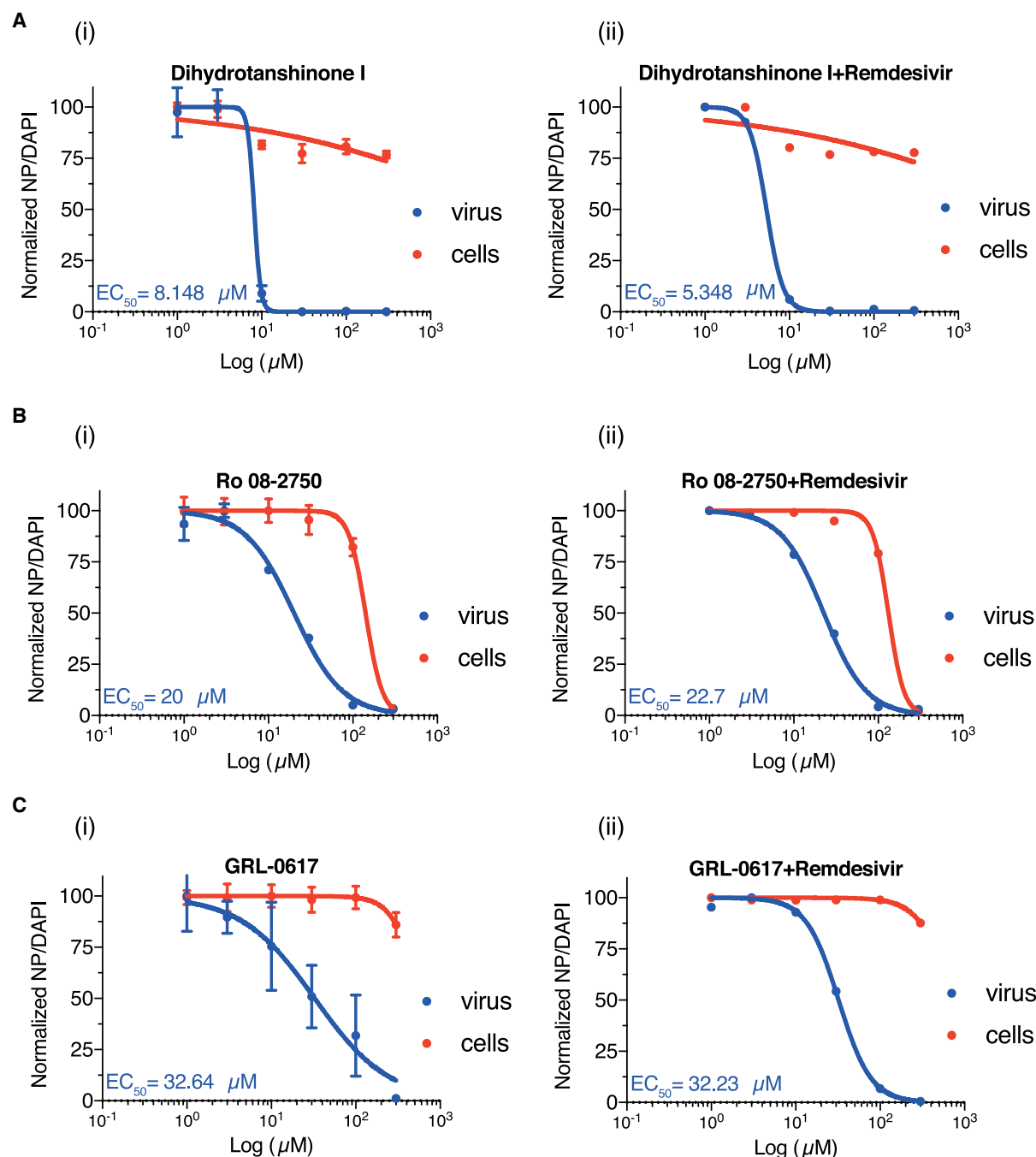


Figure 7. Determination of the EC₅₀ of dihydrotanshinone I (Ai), Ro 08-2750 (Bi), GRL-0617 (Ci) in a cell culture-based assay and the synergistic effect of each compound with 0.5 μM remdesivir (Aii, Bii and Cii).

Data was collected from three replicates.

the His-TEV- and His-Sumo-tag, respectively. After dialysis the lysate was incubated with Ni-NTA agarose beads once again to remove the proteases. The flow through was collected and loaded onto a MonoQ 5/50 GL column (GE healthcare) with buffer B, with gradient from 0.1 M to 1 M NaCl. Flow through was collected and concentrated using Amicon ultra 10 kDa (Merck). It was then loaded onto a Superdex S200 Increase 10/300 GL (GE healthcare) with buffer C (25 mM HEPES-KOH, pH 7.6, 10% glycerol, 0.02% NP-40, 150 mM NaCl and 2 mM DTT). Peak fractions were collected and pooled.

Baculovirus

3xFlag-His₆-PLpro (3FH-PLpro) was expressed in baculovirus-infected insect cells. The coding sequence was codon-optimised for *S. frugiperda* and synthesized (GeneArt, Thermo Fisher Scientific). PLpro DNA was subcloned into the biGbac vector pLIB [32] to include an N-terminal 3xFlag-His₆ tag (sequence: MDYKDHDGDYKDHDIDYKDDDDKGSHHHHHSAVLQ) (Addgene ID 169194). Baculoviruses were generated using the EMbaY baculoviral genome [33] in Sf9 cells (Thermo Fisher Scientific). For protein expression Sf9 cells were infected with baculovirus and collected 48 h after infection, flash-frozen and stored at -70°C .

PLpro gel-based assay substrate

TLKGG//APTKV (nsp2/3 junction) was used as the substrate for gel-based assay. GST-tag and MBP-tag were attached to its N- and C-terminus, respectively (Addgene ID 169195). The cleavage at G//A resulted in 25 and 42 kDa products. The constructs were expressed in T7 express lysY/I^q *E. coli* cell (NEB) as follows. Cells were grown at 37°C to log phase to achieve OD 0.8. Cells were then induced by the addition of 1 mM IPTG at 37°C for 4 h. Cells were harvested and lysed in buffer D (50 mM Tris-HCl, pH 7.5, 10% glycerol, 1 mM DTT, 0.02% NP-40, 300 mM NaCl), with the addition of 1× protease inhibitor Leupeptin (Merck), Pepstatin A (Sigma), AEBSF (Sigma), 100 μg/ml lysozyme and sonicated for 24×5 s. Lysate was centrifuged and supernatant was collected. The supernatant was incubated with amylose resin (NEB) for 2 h at 4°C . Beads were washed with wash buffer D. The protein was eluted with 10 mM maltose. Peak fractions were collected and concentrated using Amicon ultra 10 kDa (Merck). It was then loaded on Superdex S200 Increase 10/300 GL (GE healthcare) with buffer C. Peak fractions were collected and pooled.

Ulp1 catalytic fragment

An expression plasmid pFGET19-Ulp1 from Addgene. Plasmids were transformed into T7 express lysY/I^q *E. coli* cell (NEB) for expression. Cells were grown at 37°C to log phase to achieve OD 0.8. Cells were then induced by the addition of 1 mM IPTG at 37°C for 4 h. Cells were harvested and lysed within buffer E (50 mM Tris-HCl, pH 8, 5 mM magnesium acetate, 10% glycerol, 1 mM DTT, 0.02% NP-40, 500 mM NaCl, 20 mM imidazole), with the addition of 1× protease inhibitor Leupeptin (Merck), Pepstatin A (Sigma-Aldrich), AEBSF (Sigma-Aldrich), 100 μg/ml lysozyme and sonicate 24×5 s. Lysate was centrifuged and supernatant was collected. The supernatant was incubated with Ni-NTA agarose beads (Thermo Scientific) for 1 h at 4°C . Beads were washed with wash buffer D. The protein was eluted with 250 mM imidazole. Peak fractions were collected and concentrated using Amicon ultra 10 kDa (Merck). It was then loaded on Superdex S200 Increase 10/300 GL (GE healthcare) with buffer F (50 mM Tris-HCl, pH 8, 5 mM magnesium acetate, 10% glycerol, 0.5 mM TCEP, 0.01% NP-40, 500 mM NaCl). Peak fractions were collected and pooled.

Fluorescent-based assay

Three synthetic peptide substrates were used- Pro1: ELNGG//AYTRY (nsp1–2 junction), Pro2: TLKGG//APTKV (nsp2–3 junction) and Pro3: FTLKGG//APTKVT (nsp2–3 junction). Anthranilate (2-aminobenzoyl-Abz) and nitro-L-tyrosine [Y(3-NO₂)R] were attached at the N- and C-termini as fluorescent donor and quencher, respectively. The substrates were made in-house by the Peptide Synthesis department in the Francis Crick Institute. Fluorescence was read on a Spark multimode microplate reader (Tecan) with an excitation wavelength at 330 nm and emission wavelength at 420 nm. The buffer used for the assay was buffer G: 50 mM HEPES-KOH, pH 7.6, 2 mM DTT, 10% glycerol, 0.02% Tween-20, 20 μM substrate and 1.75 μM enzyme, unless it was stated in the figure.

Enzymatic assay

Kinetic constants were measured using a fluorescent-based assay. 0.5 μM enzyme were used in the titration of substrate Pro3 from 800 to 1.43 μM with 0.75-fold. V_{max} and K_m were estimated from non-linear Michaelis-Menten via linear regression of slope value.

High-throughput inhibitor screening

A total of over 5000 compounds (Sigma, Selleck, Enzo, Tocris, Calbiochem, and Symansis) were obtained from High-Throughput Screening (HTS) facility in Francis Crick Institute. The compounds were aliquoted into 48 384-well plates in two concentrations with final concentration of 1.25 μM and 3.75 μM in 20 μl reaction. An

amount of 5 μl of the enzyme was pre-incubated with compounds for 10 min before the addition of 15 μl substrate. Fluorescence was monitored at 2 min for 20 min with 3 min intervals giving a total of seven readings.

Data analysis

MATLAB was used to process data (more details in [20])

Gel-based assay

An amount of 0.5 μM of the enzyme was pre-incubated with the selected inhibitor compounds for 10 min. An amount of 0.5 μM of gel-based assay substrates were added and incubated at RT for 5 h. The buffer used for the assay was buffer G. The reactions were ran on 4–15% TGX gel (Bio-Rad, Hercules, CA) and stained with Instant blue stain (Expedeon).

Assays for deubiquitylation and deISGylation activity

Expression and purification of PLpro and the substrates K48-linked Ub3 and pro-ISG15 were as described previously [18]. PLpro was diluted in buffer G (50 mM HEPES-KOH pH 7.6, 2 mM DTT, 10% (v/v) glycerol, 0.02% (v/v) Tween-20) to 100 nM and pre-incubated with indicated concentrations of inhibitor at room temperature for 10 min. The substrates Ub3 and pro-ISG15 were diluted in buffer G to a concentration of 2 μM . An equal volume of substrate was then mixed with the inhibited PLpro to give a final concentration of 50 nM PLpro, 10 μM inhibitor and 1 μM substrate. The reaction was incubated for 1 h at 25°C and the reactions were stopped by addition of LDS sample buffer. The samples were separated on 4–12% NuPAGE Bis-Tris Gels (ThermoFisher Scientific) and stained using Pierce Silver Stain Kit (ThermoFisher Scientific).

Cell-based assay

SARS-CoV-2 production, infection and recombinant mAb production was done as described by [20].

Data Availability

All data in this paper can be found in FigShare (10779/crick.14529042).

Competing Interests

The authors declare that there are no competing interests associated with the manuscript.

Funding

This work was supported by the Francis Crick Institute, which receives its core funding from Cancer Research UK (FC001066), the UK Medical Research Council (FC001066), and the Wellcome Trust (FC001066). This work was also funded by a Wellcome Trust Senior Investigator Award (106252/Z/14/Z) to J.F.X.D. Yogesh Kulathu and Lee A. Armstrong are supported by Medical Research Council UK (MC_UU_00018/3), EMBO Young Investigator Programme (Yogesh Kulathu), ERC Starting grant (677623) (Yogesh Kulathu) and Lister research prize (Yogesh Kulathu). Berta Canal and Florian Weissmann received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement Nos 895786 and 844211, respectively. Theresa U. Zeisner received funding from the Boehringer Ingelheim Fonds.

Open Access Statement

Open access for this article was enabled by the participation of The Francis Crick Institute in an all-inclusive Read & Publish pilot with Portland Press and the Biochemical Society under a transformative agreement with JISC.

CRedit Author Contribution

John F. Diffley: Conceptualization, Supervision, Funding acquisition, Methodology, Project administration, Writing — review and editing. **Chew Theng Lim:** Conceptualization, Formal analysis, Validation, Investigation, Visualization, Methodology, Writing — original draft, Writing — review and editing. **Kang Wei Tan:** Conceptualization, Formal analysis, Validation, Investigation, Visualization, Methodology, Writing — original draft, Writing — review and editing. **Mary Wu:** Resources, Investigation, Methodology. **Rachel Ulferts:** Investigation, Methodology. **Lee A. Armstrong:** Investigation, Visualization, Writing — original draft, Writing — review and editing. **Eiko Ozono:** Validation, Investigation. **Lucy S. Drury:** Validation, Investigation, Visualization, Writing —

original draft. **Jennifer C. Milligan**: Investigation, Methodology. **Theresa U. Zeisner**: Methodology. **Jingkun Zeng**: Software. **Florian Weissmann**: Resources. **Berta Canal**: Resources. **Ganka Bineva-Todd**: Resources. **Michael Howell**: Supervision. **Nicola O'Reilly**: Resources. **Rupert Beale**: Supervision. **Yogesh Kulathu**: Supervision. **Karim Labib**: Supervision.

Acknowledgements

We thank Anne Early and Agustina P. Bertolin for their assistance and High-Throughput Screening (HTS) for dispensing the compound libraries.

Abbreviations

3CLpro, 3C-like protease; DUB, deubiquitylase; FDA, food and drug administration; FRET, fluorescence resonance energy transfer; MERS, Middle East Respiratory Syndrome; PLpro, papain-like protease; SARS-CoV-2, Severe Acute Respiratory Syndrome coronavirus 2.

References

- Dong, E., Du, H. and Gardner, L. (2020) An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect. Dis.* **20**, 533–534 [https://doi.org/10.1016/S1473-3099\(20\)30120-1](https://doi.org/10.1016/S1473-3099(20)30120-1)
- Riou, J. and Althaus, C.L. (2020) Pattern of early human-to-human transmission of Wuhan 2019 novel coronavirus (2019-nCoV), December 2019 to January 2020. *Eurosurveillance* **25**, 2000058 <https://doi.org/10.2807/1560-7917.ES.2020.25.4.2000058>
- Bernard Stoecklin, S., Rolland, P., Silue, Y., Mailles, A., Campese, C., Simondon, A. et al. (2020) First cases of coronavirus disease 2019 (COVID-19) in France: surveillance, investigations and control measures, January 2020. *Eurosurveillance* **25**, 2000094 <https://doi.org/10.2807/1560-7917.ES.2020.25.6.2000094>
- Wu, F., Zhao, S., Yu, B., Chen, Y.M., Wang, W., Song, Z.G. et al. (2020) A new coronavirus associated with human respiratory disease in China. *Nature* **579**, 265–269 <https://doi.org/10.1038/s41586-020-2008-3>
- Lu, R., Zhao, X., Li, J., Niu, P., Yang, B., Wu, H. et al. (2020) Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet* **395**, 565–574 [https://doi.org/10.1016/S0140-6736\(20\)30251-8](https://doi.org/10.1016/S0140-6736(20)30251-8)
- Matthay, M.A. and Thompson, B.T. (2020) Dexamethasone in hospitalised patients with COVID-19: addressing uncertainties. *Lancet Respir. Med.* **8**, 1170–1172 [https://doi.org/10.1016/S2213-2600\(20\)30503-8](https://doi.org/10.1016/S2213-2600(20)30503-8)
- Beigel, J.H., Tomashek, K.M., Dodd, L.E., Mehta, A.K., Zingman, B.S., Kallil, A.C. et al. (2020) Remdesivir for the treatment of covid-19: final report. *N. Engl. J. Med.* **383**, 1813–1826 <https://doi.org/10.1056/NEJMoa2007764>
- Cohen, M.S. (2021) Monoclonal antibodies to disrupt progression of early Covid-19 infection. *N. Engl. J. Med.* **384**, 289–291 <https://doi.org/10.1056/NEJMe2034495>
- Polack, F.P., Thomas, S.J., Kitchin, N., Absalon, J., Gurtman, A., Lockhart, S. et al. (2020) Safety and efficacy of the BNT162b2 mRNA covid-19 vaccine. *N. Engl. J. Med.* **383**, 2603–2615 <https://doi.org/10.1056/NEJMoa2034577>
- Voysey, M., Clemens, S.A.C., Madhi, S.A., Weckx, L.Y., Folegatti, P.M., Aley, P.K. et al. (2021) Safety and efficacy of the ChAdOx1 nCoV-19 vaccine (AZD1222) against SARS-CoV-2: an interim analysis of four randomised controlled trials in Brazil, South Africa, and the UK. *Lancet* **397**, 99–111 [https://doi.org/10.1016/S0140-6736\(20\)32661-1](https://doi.org/10.1016/S0140-6736(20)32661-1)
- Anderson, E.J., Roupael, N.G., Widge, A.T., Jackson, L.A., Roberts, P.C., Makhene, M. et al. (2020) Safety and immunogenicity of SARS-CoV-2 mRNA-1273 vaccine in older adults. *N. Engl. J. Med.* **383**, 2427–2438 <https://doi.org/10.1056/NEJMoa2028436>
- Kirby, T. (2021) New variant of SARS-CoV-2 in UK causes surge of COVID-19. *Lancet Respir. Med.* **9**, e20–e21 [https://doi.org/10.1016/S2213-2600\(21\)00005-9](https://doi.org/10.1016/S2213-2600(21)00005-9)
- Maier, H.J., Bickerton, E. and Britton, P. (2015) Coronaviruses: methods and protocols. *Methods Mol. Biol.* **1282**, 1–282 https://doi.org/10.1007/978-1-4939-2438-7_1
- Báez-Santos, Y.M., Barraza, S.J., Wilson, M.W., Agius, M.P., Mielech, A.M., Davis, N.M. et al. (2014) X-ray structural and biological evaluation of a series of potent and highly selective inhibitors of human coronavirus papain-like proteases. *J. Med. Chem.* **57**, 2393–2412 <https://doi.org/10.1021/jm401712t>
- Báez-Santos, Y.M., St. John, S.E. and Mesecar, A.D. (2015) The SARS-coronavirus papain-like protease: Structure, function and inhibition by designed antiviral compounds. *Antiviral Res.* **115**, 21–38 <https://doi.org/10.1016/j.antiviral.2014.12.015>
- Lei, J., Kusov, Y. and Hilgenfeld, R. (2018) Nsp3 of coronaviruses: Structures and functions of a large multi-domain protein. *Antiviral Res.* **149**, 58–74 <https://doi.org/10.1016/j.antiviral.2017.11.001>
- Blanchard, S.C., Kim, H.D., Gonzalez, R.L., Puglisi, J.D. and Chu, S. (2004) tRNA dynamics on the ribosome during translation. *Proc. Natl Acad. Sci. U.S.A.* **101**, 12893–12898 <https://doi.org/10.1073/pnas.0403884101>
- Armstrong, L.A., Lange, S.M., de Cesare, V., Matthews, S.P., Nirujogi, R.S., Cole, I. et al. (2020) Characterization of protease activity of Nsp3 from SARS-CoV-2 and its inhibition by nanobodies. *bioRxiv* <https://doi.org/10.1101/2020.12.09.417741>
- Han, Y.-S., Chang, G.-G., Juo, C.-G., Lee, H.-J., Yeh, S.-H., Hsu, J.T.-A. et al. (2005) Papain-like protease 2 (PLP2) from severe acute respiratory syndrome coronavirus (SARS-CoV): expression, purification, characterization, and inhibition. *Biochemistry* **44**, 10349–10359 <https://doi.org/10.1021/bi0504761>
- Zeng, J., Weissmann, F., Bertolin, A.P., Posse, V., Canal, B., Ulferts, R. et al. (2021) Identifying SARS-CoV-2 antiviral compounds by screening for small molecule inhibitors of nsp13 helicase. *Biochem. J.* **478**, 2405–2423 <https://doi.org/10.1042/BCJ20210201>
- Park, J.Y., Kim, J.H., Kim, Y.M., Jeong, H.J., Kim, D.W., Park, K.H. et al. (2012) Tanshinones as selective and slow-binding inhibitors for SARS-CoV cysteine proteases. *Bioorg. Med. Chem.* **20**, 5928–5935 <https://doi.org/10.1016/j.bmc.2012.07.038>

- 22 Freitas, B.T., Durie, I.A., Murray, J., Longo, J.E., Miller, H.C., Crich, D. et al. (2020) Characterization and noncovalent inhibition of the deubiquitinase and deISGylase activity of SARS-CoV-2 papain-like protease. *ACS Infect. Dis.* **6**, 2099–2109 <https://doi.org/10.1021/acinfecdis.0c00168>
- 23 Ratia, K., Pegan, S., Takayama, J., Sleeman, K., Coughlin, M., Baliji, S. et al. (2008) A noncovalent class of papain-like protease/deubiquitinase inhibitors blocks SARS virus replication. *Proc. Natl Acad. Sci. U.S.A.* **105**, 16119–16124 <https://doi.org/10.1073/pnas.0805240105>
- 24 Shin, D., Mukherjee, R., Grewe, D., Bojkova, D., Baek, K., Bhattacharya, A. et al. (2020) Papain-like protease regulates SARS-CoV-2 viral spread and innate immunity. *Nature* **587**, 657–662 <https://doi.org/10.1038/s41586-020-2601-5>
- 25 Basu, S., Mak, T., Ulferts, R., Wu, M., Deegan, T., Fujisawa, R. et al. (2021) Identifying SARS-CoV-2 antiviral compounds by screening for small molecule inhibitors of Nsp14 RNA cap methyltransferase. *Biochem. J.* **478**, 2481–2497 <https://doi.org/10.1042/BCJ20210219>
- 26 Li, M., Li, Q., Zhang, C., Zhang, N., Cui, Z., Huang, L. et al. (2013) An ethnopharmacological investigation of medicinal *Salvia* plants (Lamiaceae) in China. *Acta Pharm. Sin. B* **3**, 273–280 <https://doi.org/10.1016/j.apsb.2013.06.001>
- 27 Klemm, T., Ebert, G., Calleja, D.J., Allison, C.C., Richardson, L.W., Bernardini, J.P. et al. (2020) Mechanism and inhibition of the papain-like protease, PLpro, of SARS-CoV-2. *EMBO J.* **39**, e106275 <https://doi.org/10.15252/embj.2020106275>
- 28 Smith, E., Davis-Gardner, M.E., Garcia-Ordonez, R.D., Nguyen, T.-T., Hull, M., Chen, E. et al. (2020) High-throughput screening for drugs that inhibit papain-like protease in SARS-CoV-2. *SLAS Discov.* **25**, 1152–1161 <https://doi.org/10.1177/2472555220963667>
- 29 Diniz, L.R.L., Perez-Castillo, Y., Elshabrawy, H.A., Filho, C.D.S.M.B. and de Sousa, D.P. (2021) Bioactive terpenes and their derivatives as potential SARS-CoV-2 proteases inhibitors from molecular modeling studies. *Biomolecules* **11**, 74 <https://doi.org/10.3390/biom11010074>
- 30 Mielech, A.M., Chen, Y., Mesecar, A.D. and Baker, S.C. (2014) Nidovirus papain-like proteases: multifunctional enzymes with protease, deubiquitinating and deISGylating activities. *Virus Res.* **194**, 184–190 <https://doi.org/10.1016/j.virusres.2014.01.025>
- 31 Rambaut, A., Loman, N., Pybus, O., Barclay, W., Barrett, J., Carabelli, A. et al. (2020) Preliminary genomic characterisation of an emergent SARS-CoV-2 lineage in the UK defined by a novel set of spike mutations/563 <https://virological.org/t/preliminary-genomic-characterisation-of-an-emergent-sars-cov-2-lineage-in-the-uk-defined-by-a-novel-set-of-spike-mutations/563>
- 32 Weissmann, F., Petzold, G., VanderLinden, R., Huis In 't Veld, P.J., Brown, N.G., Lampert, F. et al. (2016) biGBac enables rapid gene assembly for the expression of large multisubunit protein complexes. *Proc. Natl Acad. Sci. U.S.A.* **113**, E2564–E2569 <https://doi.org/10.1073/pnas.1604935113>
- 33 Trowitzsch, S., Bieniossek, C., Nie, Y., Garzoni, F. and Berger, I. (2010) New baculovirus expression tools for recombinant protein complex production. *J. Struct. Biol.* **172**, 45–54 <https://doi.org/10.1016/j.jsb.2010.02.010>

Supplemental Materials

Figure S1. (Related to Figure 1) Purification of PLpro in bacteria and baculovirus.

- A. Purification of His-TEV-PLpro from bacteria. Lanes 11-13 contain monomer of PLpro from the last step of the purification on a gel filtration column and these were pooled.
- B. Purification of Flag-His-PLpro from baculovirus. Lane 2-4 contain monomer of PLpro from a gel filtration column and were pooled. The tag remains intact with protein.

Figure S2. (Related to Figure 2) Optimization of the FRET assay for PLpro.

- A. PLpro was able to cleave substrate Pro2 (nsp2/3) but not Pro1 (nsp1/2) (see Material & Methods)
- B. PLpro was able to cleave substrate Pro3 (12 amino acids) more efficiently than Pro2 (10 amino acids).
- C. Comparison between the TEV-tagged and Sumo-tagged versions of PLpro enzyme activities.
- D. Comparison between the baculovirus and bacterial (TEV-tagged) purified PLpro.

Figure S3. (Related to Figure 4) Gel-based assay.

- A. Purification of the substrate (67 kDa) with Superdex S200. Fractions 14-18 were pooled and used for the gel-based assay.
- B. Dihydrotanshinone I, beta-lapachone, cryptotanshinone, tanshinone IIA, Cdk4 inhibitor III, GLR-0617, and Ro 08-2750 did not inhibit 3CLpro (nsp5) activity. Z-VAD-FMK is a novel inhibitor discovered against nsp5 (reference nsp5 paper in this issue)

Figure S4. EC₅₀ of beta-lapachone (A), Cdk4 inhibitor III (B), cryptotanshinone (C), and tanshinone IIA (D) from the cell culture-based viral proliferation assay.

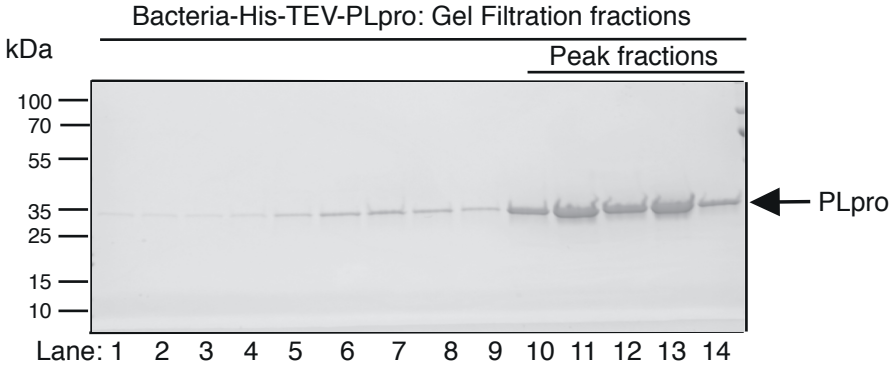
Figure S5. (Related to Figure 7) (A and C) Viral proliferation assay for dihydrotanshinone I, Ro 08-2750, GRL-0617, beta-lapachone, cryptotanshinone, tanshinone IIA, and Cdk4 inhibitor III. B. There is no obvious additional effect of

combining remdesivir with compounds in (A). These compounds either do not stop COVID-19 viral proliferation or are cytotoxic.

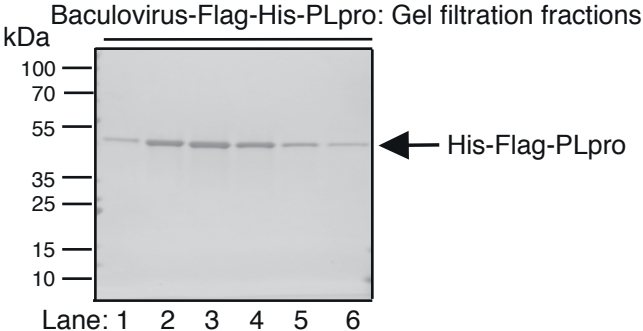
Table S1. Compound list of over 5000 compounds with scores from low and high concentration.

Table S2. Oligomers and PLpro optimized sequence.

A

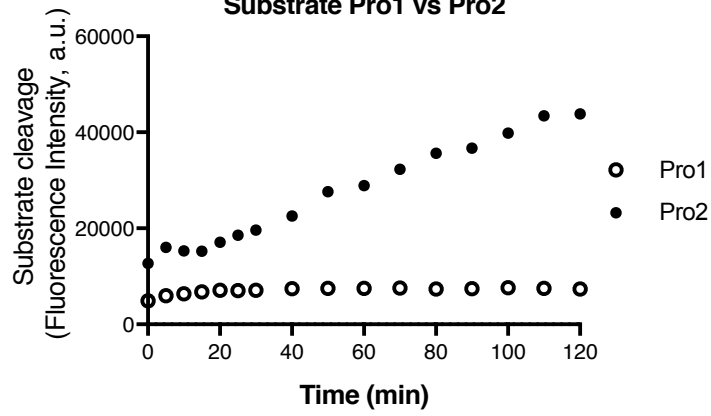


B



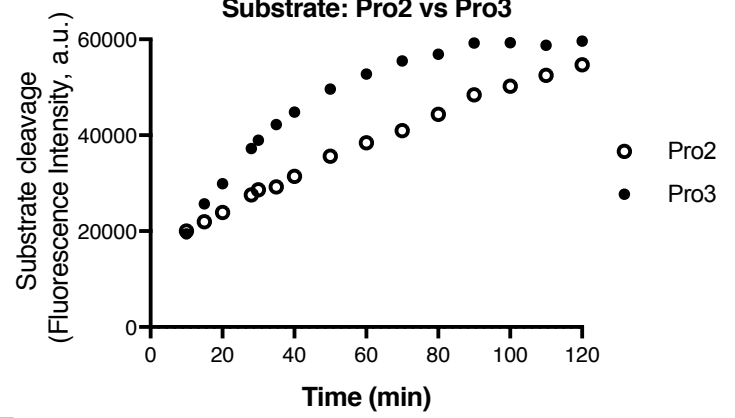
A

Substrate Pro1 vs Pro2



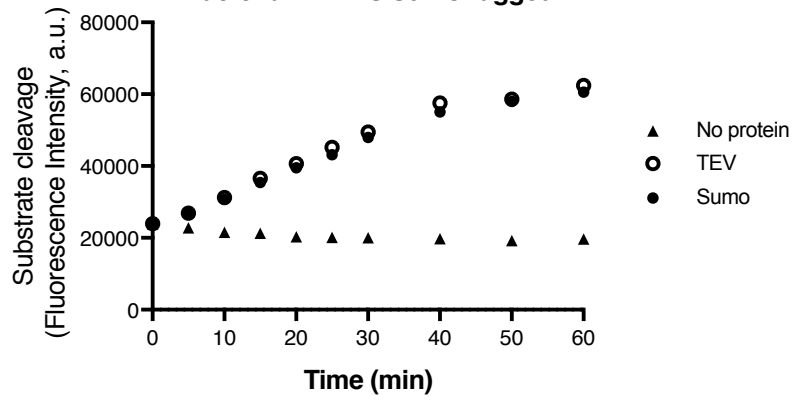
B

Substrate: Pro2 vs Pro3



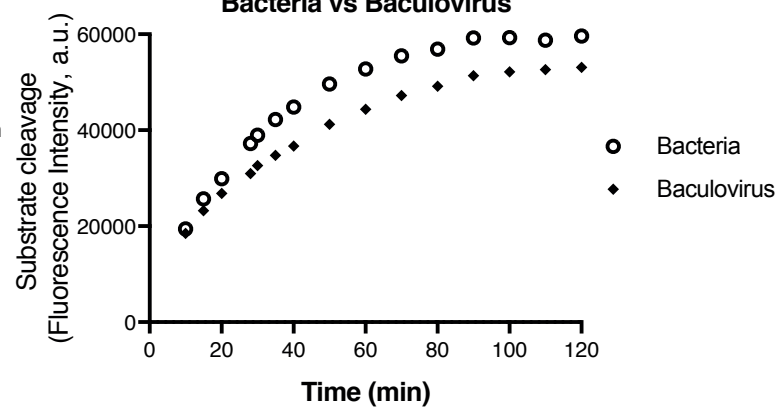
C

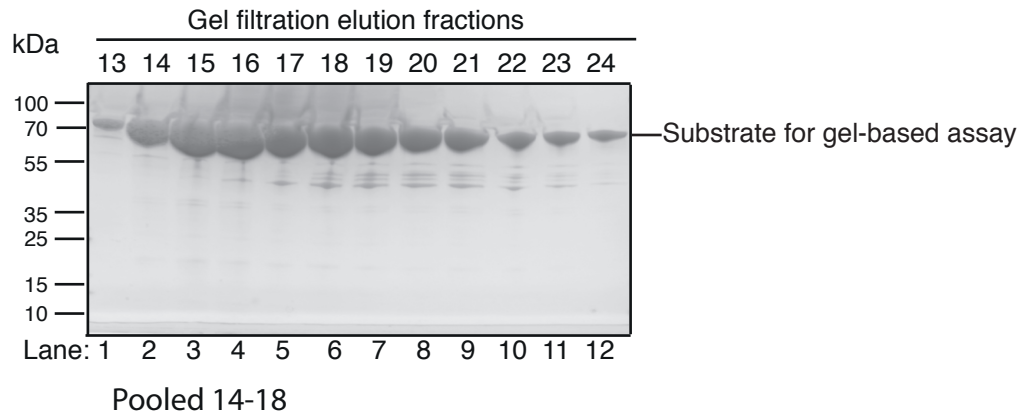
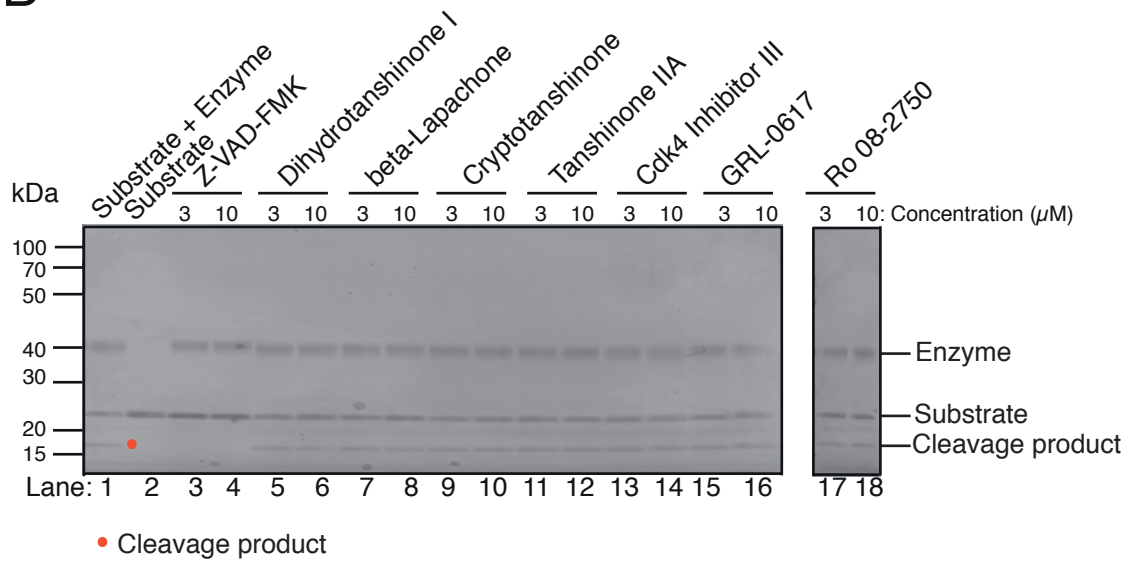
Bacteria TEV- vs Sumo-tagged

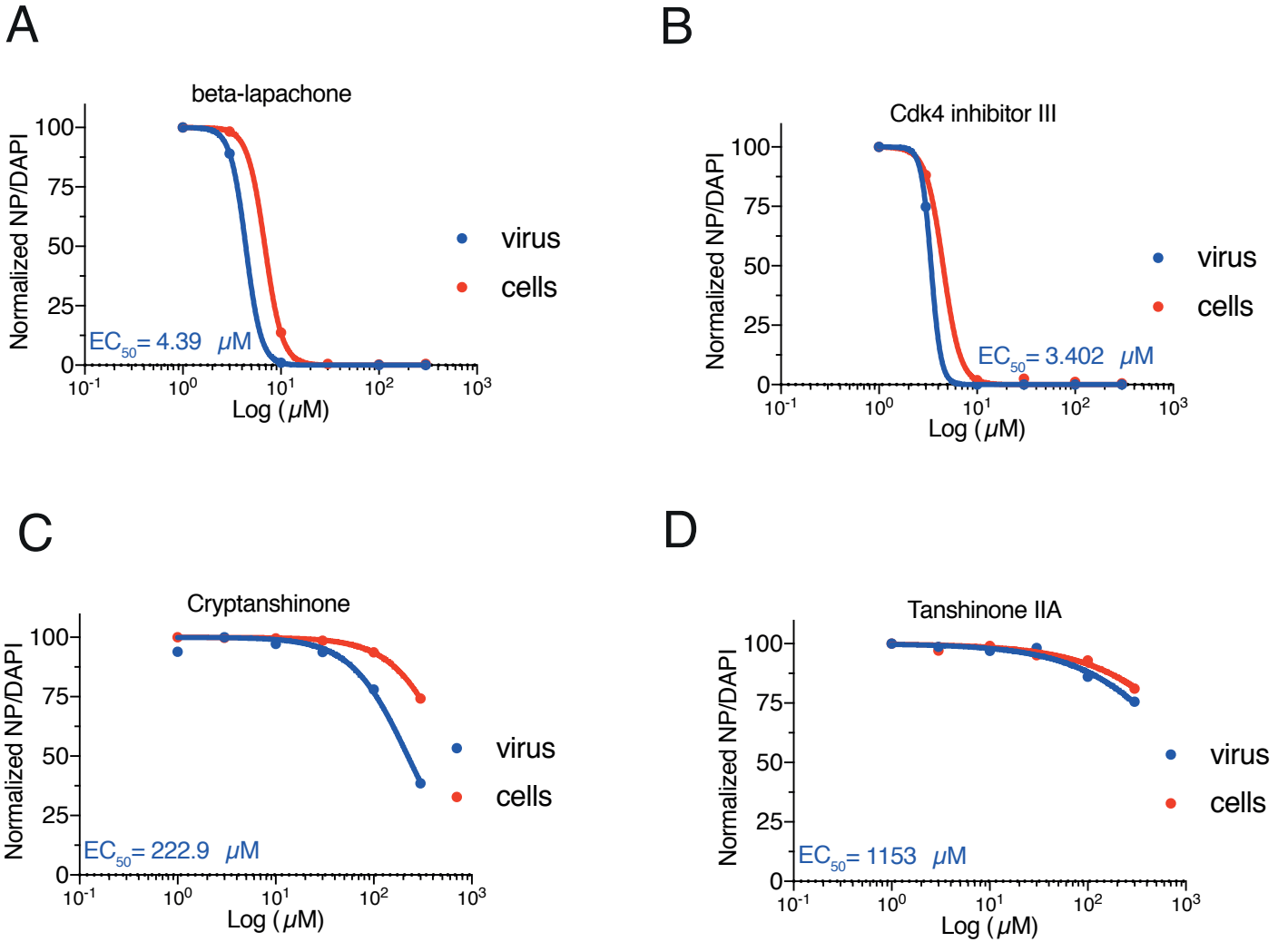


D

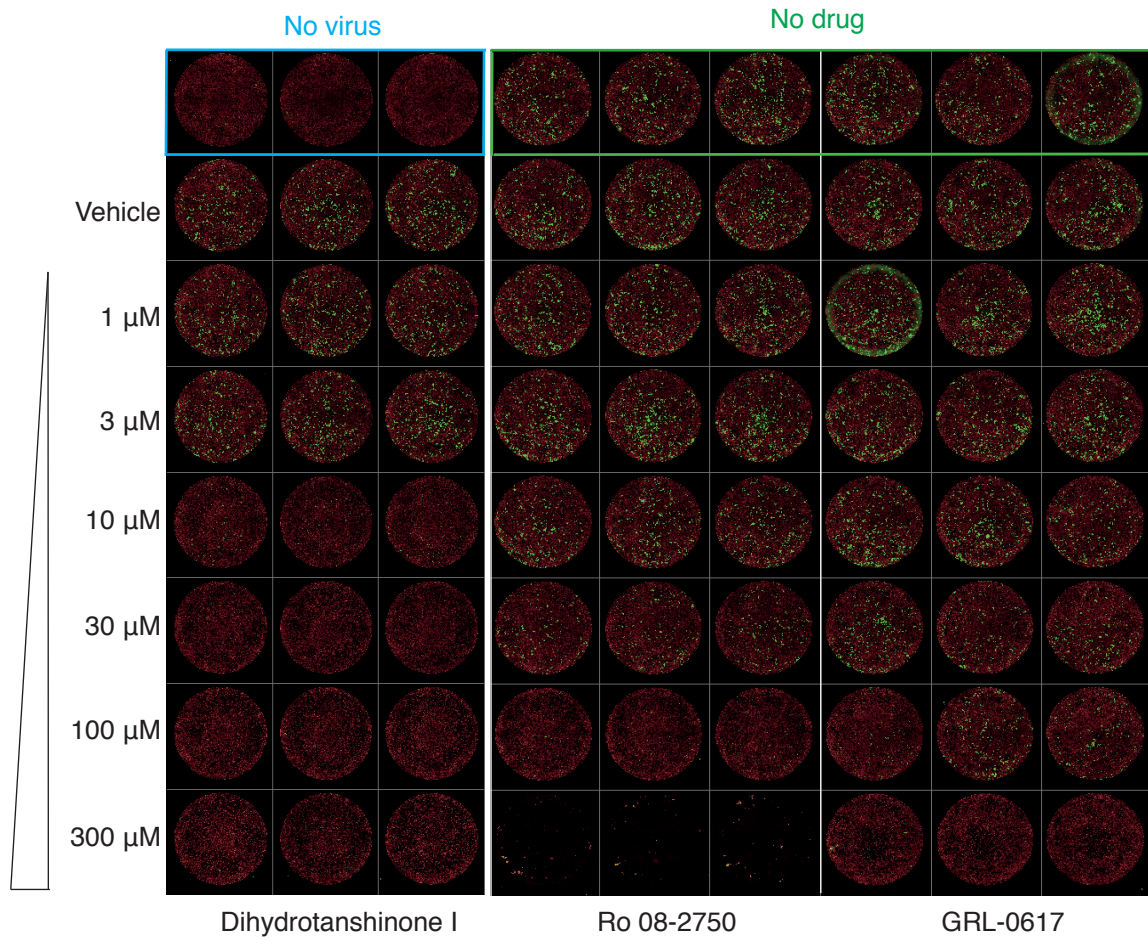
Bacteria vs Baculovirus



A**B**



A



B

