Main protease structure and XChem fragment screen

Summary

To contribute to the global effort to combat COVID-19, Diamond has been able to solve a new structure of the SARS-CoV-2 main protease (MPro) at high resolution (PDB ID: 6YB7), and complete a large XChem crystallographic fragment screen against it (detailed below). Data have been deposited with the PDB, but we are making the results available immediately to the world on this page; additional work is ongoing, and updates will be continually posted here in coming days and weeks.

This work builds on the sensationally fast crystal structure of MPro at 2.16 Å in complex with a covalent inhibitor, released in January this year by Prof Zihe Rao (6LU7, published here, described here). We thus ordered the synthetic gene and cloned the full length protein as previously described for the SARS main protease (Xue et al 2007). This yielded crystals of the unliganded enzyme that diffracted to high resolution (1.25 Å) on beamline I04-1, in a different space group to the inhibitor complex, and the structure was determined and refined rapidly. Critically, this showed it had the active site empty and solvent accessible - perfect for fragment screening.

So it proved: the first 600-crystal experiment could be completed in 72 hours, through growing large numbers of crystals, optimising the soaking conditions, soaking and harvesting all 600 crystals and completing the data collection run on beamline I04-1. The hits from this initial run and other details were pre-released on March 6th.

By 24th March 2020, the initial 1500-crystal experiment was complete, and the results made publicly available. Screening additional libraries throughout April brought the total number of active site fragments to 71, with 48 fragments binding covalently. This was an exceptionally large screen which yielded a remarkably rich readout, with vast opportunities for fragment growing and merging.

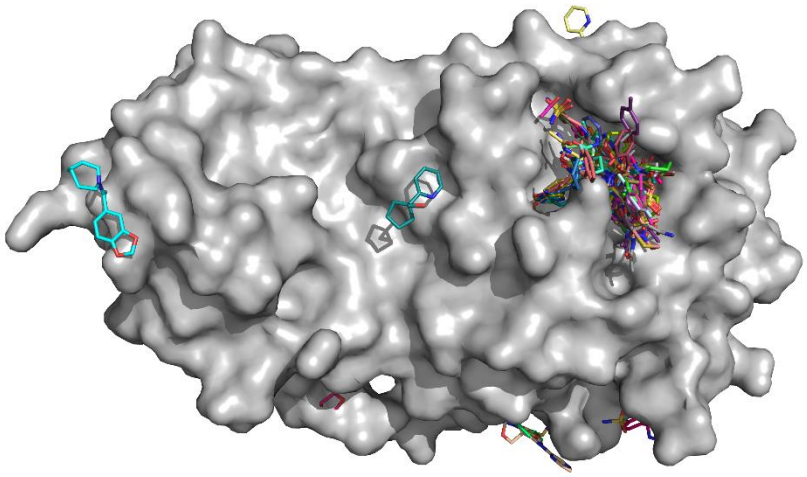
We have already triggered computationally-driven follow-up work internally, and externally joined forces to launch a fully-open crowdsourcing and crowdfunding initiative – the COVID Moonshot - to establish urgently the shortest route possible to clinical impact by maximally exploiting the readout - you can help, read more here.

On 11th May 2020, the first biochemical and structural data from Moonshot compounds was released and by 12th June, over 500 compounds had been tested, demonstrating that the design-make-test process is fully in place.

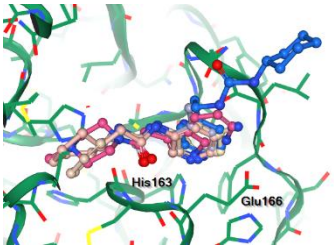
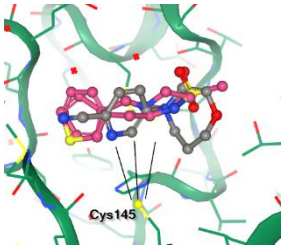
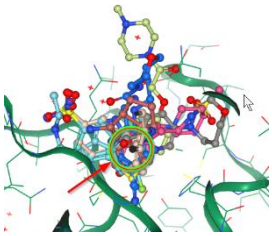
XChem fragment screen

The initial screen encompassed multiple fragment libraries: the DSI-poised library, MiniFrags (Astex) FragLites and Peplites (CRUK Newcastle Drug Discovery Unit (Newcastle University)), York3D (University of York), SpotFinder and heterocyclic electrophilic fragment library (Hungarian Academy of Sciences) and an electrophilic fragment library designed and pre-screened by mass spec at the Weizmann Institute (see below).
There were 74 hits of high interest - data and extensive details are here, and some interactive views here:

- 23 non-covalent hits in the active site
- 48 covalent hits in the active site
- 3 hits in the dimer interface, one in a calculated hotspot

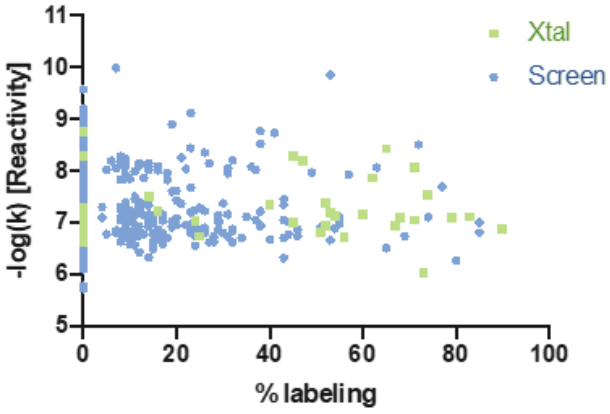


Features of note include:

S1 - the His163 - Glu166 motif	S2 - hits over catalytic Cys145	S3 - the aromatic wheel
		

Mass spectrometry screening of electrophiles

The London lab screened M^{pro} supplied by Diamond with 993 fragment electrophiles by their intact-protein mass-spectrometry platform (Resnick *et al*. JACS 2019). A preliminary screen (200 µM, 24 h, 4°C) revealed 130 compounds that irreversibly label the protein (>50%); a more stringent screen (5 µM, 1.5 h, RT) discriminated several compound series that still label robustly:

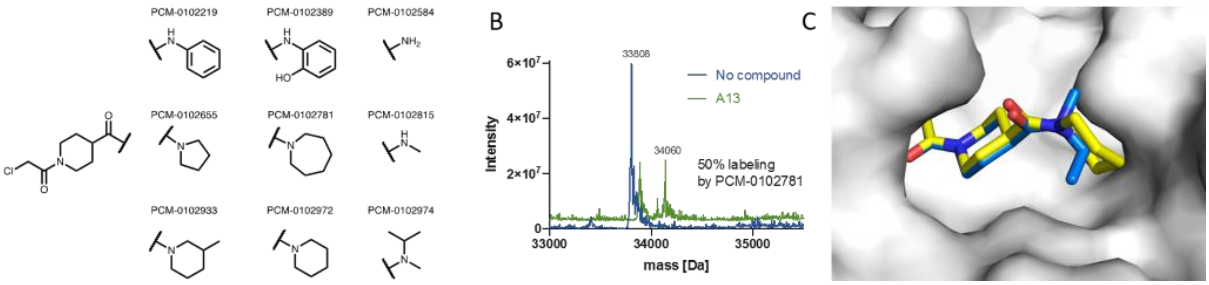


%labeling of fragment electrophile (x-axis) vs. their intrinsic reactivity (y-axis), with lower reactivity lying higher up the axis. Most interesting are

low reactivity compounds that achieve high labeling, which point to selective and thus low-toxicity compounds.

The full experiment is summarised in this spreadsheet: 74 compounds were screened at XChem by crystal soaking and co-crystallisation following Resnick *et al*, working in triplicate: resulting in 42 complex crystal structures. (Of the rest, 29 yielded crystals without bound ligand, and 3 could not be measured for experimental reasons.)

A particularly promising series are the chloroacetamides based on piperidine-amides: as they were not hits in many previous screens, they should be very selective to M^{pro}. Additionally, they reveal SAR (Structure activity relationships) even amongst the nine representatives in the library, suggesting binding can be optimised. Based on the several crystal structures of these hits, we have commenced synthesis of analogs and testing of optimised derivatives.

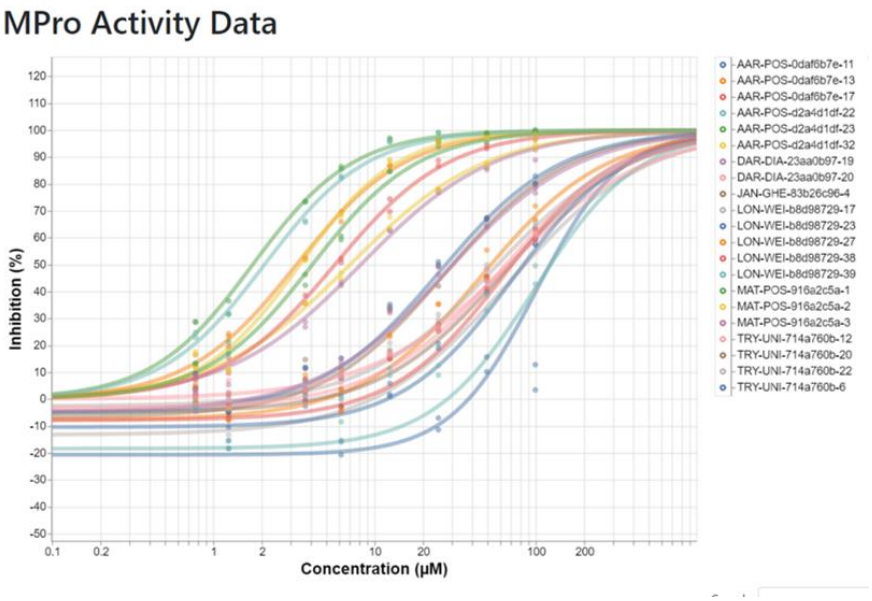


Structure activity relationship of a selective hit series. A. The piperidine-amides that showed robust labeling at 5 µM, 1.5 h RT incubation: various side-chains off of the amide show varying degrees of labeling, the aniline substitution especially potent. **B.** Example MS spectrum of M^{pro} protease alone (blue) or incubated with five covalent fragments (5 µM, 1.5 h, RT); only one showed labeling at ~50%. **C.** Crystal structures of PCM-0102972 (yellow) and PCM-0102974 (blue) showing optimisation vectors.

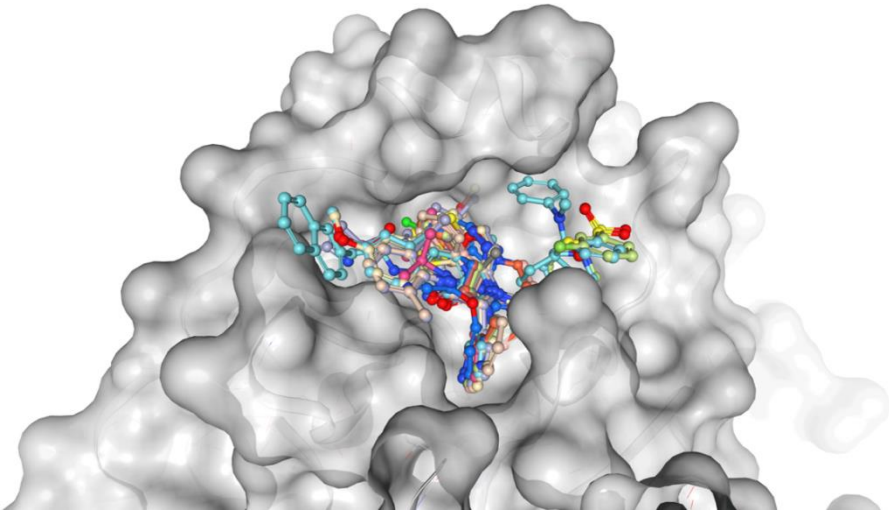
COVID MoonShot - Taking fragments to impact

The very large M^{pro} screening experiment yielded an exceptionally rich set of information, with extensive opportunities for fragment growing and merging; we are thus anxious to ensure the high quality data is as effective as possible at seeding anti-COVID-19 efforts around the world.

We have joined forces in a fully open initiative - the COVID MoonShot - to drive these results as far towards clinical impact as possible, by **crowdsourcing molecule design** from medicinal chemists worldwide and converting these into physical compounds and assays.
To date (16th June 2020), over 500 MoonShot compounds have been screened in the XChem lab, in a fluorescence-based assay by our collaborators at the Weizmann Institute (Israel), as well as in a mass-spec-based assay run by the Schofield group (CRL, University of Oxford).



Additionally, on the 11th of May, the first wave of follow-up structures was released. These, and all our other Mpro structures, can be viewed online in our Fragalysis viewer.



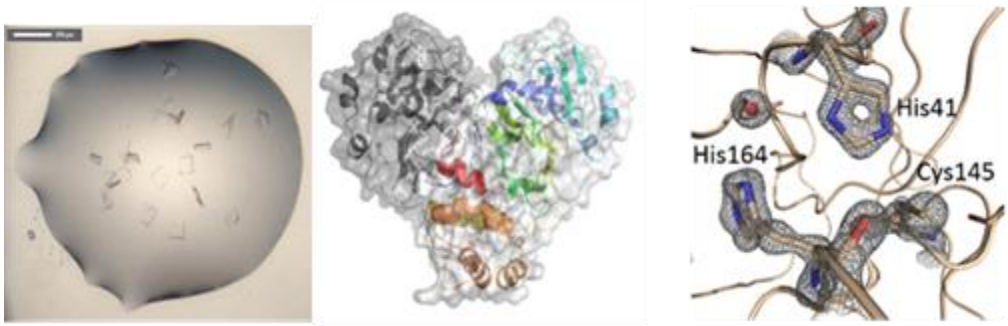
For continuous updates on assay results and structure release please consult CCD Vault or the MoonShot webpage.

Highlights on progress

Highlights achieved to date for structure-based lead discovery on the SARS-CoV-2 Main protease

- 2020-02-14 SARS-CoV-2 Main protease (MPRO) cloned and overexpressed in high yields for structure-based drug discovery (>40 mg purified protein from 1 L media)
- 2020-02-20 Crystal structure of free MPRO solved to 1.9 Å.
- 2020-02-25 Further optimisation of crystals leading to structure at 1.39 Å resolution (PDB ID 6Y84, released on 2020-03-11). These data complement a structure in complex with a covalent inhibitor (N3) determined by Z. Rao Group (Shanghai Tech) at 2.16 Å resolution. (Data were released in PDB on 2020-02-05 pdbid 6LU7)

- 2020-02-25 Using purified protein sent to the Weizmann Institute, a mass spec screen of electrophilic fragments yielded 130 preliminary hits, with 30 hits validated by a secondary stringent screen. Synthesis of analogues of high priority hits commences.
- 2020-02-25 Crystallisation conditions optimised for a full XChem fragment screening experiment, thanks to ~17,000 crystallisation experiments facilitated by the Diamond/RCaH/RFI HTP crystallisation facility at Harwell.
- 2020-02-26 Significant XChem fragment screen completed (~600 fragments tested); a preliminary analysis reveals 5 confident hits in the active site, with several more sites of interest around the protein.
- 2020-03-05 Team reach 1,000 crystals harvested before lunch… reaching 1495 crystals by end of the day.
- 2020-03-06
- 6am: the final crystal measured on beamline I04-1, which ran without hiccup through the entire experiment
- 5pm: structures of the first 7 fragment hits in active site (from first 600 fragments) released via Diamond's website.
- 2020-03-12: First release of set of ~80 hit structures fully modelled and refined (3 crystallographers working full steam all week), ready for final proofreading and preparation for release.
- 2020-03-18: 78 hits released to the public: 58 in the active site, 39 of which are covalently bound.
- 2020-03-24: 13 additional structures released taking the total to 66 active site fragments, 44 of which are covalently bound.
- 2020-04-16: Screening additional fragment libraries identifys 2 new structures with nitrile moeity acting as covalent warheads.
- 2020-05-05: First structures from MoonShot collaboration uploaded to Fragalysis.
- 2020-05-11: First biochemical results released.
- 2020-05-11: Over 500 MoonShot compounds tested.



Structure of SARS-CoV-2. (a) Crystals of SARS-CoV-2 free enzyme (b) cartoon representation of the SARS-CoV-2 dimer with a semi-transparent surface in grey (c) representative electron density (2Fo-Fc map contoured at the 2.5 s level) from the 1.39 Å structure centered on active site residues His41, His164 and Cis145

This has been a serious team effort:

Group of Martin Walsh:

- Claire Strain-Damerell - *molecular biology*

- David Owen - *crystallogenesis and crystallography*
- Petra Lukacik - *protein purification and team lead*

XChem Lab Team (Frank von Delft)

- Alice Douangamath - *XChem experiment and analysis*
- Daren Fearon - *XChem experiment and analysis*
- Anthony Aimon - *compound management and logistical support*

XChem Industrial Liaison Group

- Ailsa Powell - *XChem experiment and analysis*
- Alexandre Dias - *logistical support for the experiment*

Group of Nir London (Weizmann Institute, Israel)

- Efrat Resnick - *mass-spec screening*
- Paul Gehrtz - *synthetic chemistry*
- Rambabu Reddi - *synthetic chemistry*

Protein Crystallography group of SGC-Oxford (Frank von Delft)

- Conor Wild - *PanDDA-2*
- Tobias Krojer - *data processing and structure deposition*

Fragalysis team (Frank von Delft)

- Rachael Skyner - *data curation*
- Anna Carbery - *back-end development*

I04-1 beamline team (Frank von Delft)

- Jose Brandao-Neto
- Louise Dunnett

Diamond MX group

- Mark Williams
- David Aragao
- Adam Crawshaw
- Marco Mazzorana
- Katherine McAuley
- Ralf Flaig
- Dave Hall
- Dave Stuart

Advice and encouragement from the group of Zihe Rao (ShanghaiTech) as well as Mike Waring and Martin Noble (CRUK Newcastle Drug Discovery Unit) and György M. Keserű (Hungarian Academy of Sciences).