Discovery of Antiviral Cyclic Peptides Targeting the Main Protease of SARS-CoV-2 *via* mRNA Display

Jason Johansen-Leete,^{1,2} Sven Ullrich,³ Sarah E. Fry,^{1,2} Rebecca Frkic,^{3,4} Max J. Bedding,^{1,2} Anupriya Aggarwal,⁵ Anneliese S. Ashhurst,^{1,2,6} Kasuni B. Ekanayake,^{3,4} Mithun C. Mahawaththa,^{3,4} Vishnu M. Sasi,^{3,4} Toby Passioura,^{1,2,7,8} Mark Larence,^{8,9} Gottfried Otting,^{3,4} Stuart Turville,⁵ Colin J. Jackson,^{3,4} Christoph Nitsche,^{3*} Richard J. Payne^{1,2*}

¹ School of Chemistry, The University of Sydney, Sydney, New South Wales 2006, Australia

² Australian Research Council Centre of Excellence for Innovations in Peptide and Protein Science, The University of Sydney, Sydney, New South Wales 2006, Australia

³ Research School of Chemistry, Australian National University, Canberra, ACT 2601, Australia

⁴ Australian Research Council Centre of Excellence for Innovations in Peptide and Protein Science, Australian National University, Canberra, ACT 2601, Australia

⁵ Kirby Institute, Sydney, NSW 2052, Australia

⁶ School of Medical Sciences, Faculty of Medicine and Health, The University of Sydney, Sydney, New South Wales 2006, Australia.

⁷ School of Life and Environmental Sciences, The University of Sydney, Sydney NSW 2006, Australia

⁸ Sydney Analytical, The University of Sydney, Sydney NSW 2006, Australia

⁹ Charles Perkins Centre, The University of Sydney, Sydney, New South Wales 2006, Australia

Abstract

Antivirals that specifically target SARS-CoV-2 are needed to control the COVID-19 pandemic. The main protease (M^{pro}) is essential for SARS-CoV-2 replication and is an attractive target for antiviral development. Here we report the use of the Random nonstandard Peptide Integrated Discovery (RaPID) mRNA display on a chemically cross-linked SARS-CoV-2 M^{pro} dimer, which yielded several high-affinity thioether-linked cyclic peptide inhibitors of the protease. Structural analysis of M^{pro} complexed with a selenoether analogue of the highest-affinity peptide revealed key binding interactions, including glutamine and leucine residues in sites S₁ and S₂, respectively, and a binding epitope straddling both protein chains in the physiological dimer. Several of these M^{pro} peptide inhibitors possessed antiviral activity against SARS-CoV-2 *in*

vitro with EC₅₀ values in the low micromolar range. These cyclic peptides serve as a foundation for the development of much needed antivirals that specifically target SARS-CoV-2.

Introduction

The COVID-19 pandemic, caused by infection with severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has caused widespread morbidity and mortality as well as devastation to the global economy since the disease was first reported in late 2019 in Wuhan, China¹. At the time of writing there has been more than 200 million confirmed cases and 4.3 million deaths worldwide as a result of COVID-19². There has been significant effort from the global research community to develop effective vaccines for COVID-19; this has been enormously successful, with adenoviral vectored vaccines, protein vaccines and mRNA vaccines now in widespread use across the world. Whilst vaccines will enable protective immunity in most, there will be populations where vaccine-based immunity may fail, and these individuals will be vulnerable to SARS-CoV-2 infection in the future. It is furthermore unclear what changes will appear in the virus in contemporary SARS-CoV-2 viral variants (highlighted by the recent emergence of the delta variant B.1.617.2)³, and how those variants will navigate both convalescent and vaccine immune responses. Given that this is the third coronavirus that has crossed via zoonoses, antiviral development against SARS-CoV-2, and future coronaviruses with pandemic potential, are desperately needed in addition to prophylactic vaccines.

While there have been significant efforts toward the discovery of effective antivirals for SARS-CoV-2, the vast majority of molecules that have completed (or are currently

being assessed in) clinical trials were originally developed for other infectious and inflammatory disease indications and are being repurposed for COVID-19. For example, remdesivir, originally trialled for Ebola, is currently the only antiviral drug to be approved by the U.S. Food and Drug Administration (FDA) for the treatment of COVID-19. While the molecule has been shown to possess some activity during early infection, it has shown limited to no efficacy in a number of trials^{4,5}, as well as in infections in patients hospitalized with COVID-19⁶. Other repurposed antiviral drugs that have entered trials include the HIV combination therapy lopinavir-ritonavir^{7,8}, type I interferon treatments^{9,10}, and the antimalarial hydroxychloroquine¹¹⁻¹³; however, these have not demonstrated improvement in disease progression over standard care. In fact, it has recently been suggested that many repurposing efforts may be compromised by experimental artefacts reflecting the physicochemical properties of certain drugs rather than specific target-based activities¹⁴. To date, the most effective therapeutic intervention for improving COVID-19 patient outcomes in a hospital setting is the use of the corticosteroid dexamethasone, which reduces inflammation-mediated lung injury associated with SARS-CoV-2 infection in patients with elevated levels of C-reactive protein^{15 16}. Based on the above, there is an urgent need for the discovery of effective antivirals for COVID-19, ideally with mechanisms of action that specifically target proteins critical in the SARS-CoV-2 lifecycle.

Infection of human cells by SARS-CoV-2 is initiated by interaction between the receptor binding domain of the trimeric viral spike protein (S) with the host cell-surface receptor angiotensin converting enzyme 2 (ACE2) (Figure 1). Following receptor binding of the virus, the spike protein is activated by cleavage between the S1 and S2 domains leading to host cell entry via two distinct pathways: 1) an endocytic pathway

through endosomal–lysosomal compartments with spike cleavage facilitated by lysosomal cathepsins, or 2) a cell surface pathway following activation by a serine protease such as TMPRSS2^{17–19}. Following proteolysis, the N-terminus of the cleaved S2 domain is embedded into the cell membrane and leads to fusion of the membranes of the virus and the host cell, followed by transfer of viral RNA into the cytoplasm²⁰.

Fig. 1. Mechanism of SARS-CoV-2 entry into host cells and replication (cell-surface entry pathway mediated by TMPRSS2 shown). The proteolytic activity of SARS-CoV-2 M^{pro} on the two polyproteins pp1a and pp1ab is shown in the box, including the structure of the covalent peptidomimetic M^{pro} inhibitor PF-07321332 developed by Pfizer.

Viral gene expression within the host cell results in the translation of two overlapping polyproteins, pp1a and pp1ab. Embedded within these polyproteins are sixteen non-

structural proteins critical for viral replication, the majority of which form the viral replication and transcription complex (RTC) including the viral RNA-dependent RNA polymerase (RdRp, nsp12) and helicase (nsp13)²⁰. These proteins become functional only after proteolytic release by two viral proteases. The first of these is a domain of nsp3 called the papain-like protease (PL^{pro}) which cleaves the pp1a and pp1ab at three sites, releasing nsp1, nsp2 and nsp3^{20,21}. The second is the SARS-CoV-2 main protease (M^{pro}), also called nsp5 or the chymotrypsin-like protease (3CL^{pro}), which cleaves pp1a and pp1ab at a minimum of 11 distinct cleavage sites to release nsp4-16 (Figure 1)²⁰. Interestingly, M^{pro} has also been found to aid in immune evasion by inhibiting type I IFN production, contributing to the impaired type I IFN response that has become a hallmark of severe SARS-CoV-2 infection, with persistent viral load and poor patient outcomes²²⁻²⁵. M^{pro} forms a catalytically active homodimer which cleaves with high specificity at Leu-Gln↓Xaa (where ↓ represents the cleavage site and Xaa can be Ser, Ala or Asn)²⁶⁻²⁸. Such sequence specificity has not been observed for any human proteases and therefore peptide or peptidomimetic based inhibitors are predicted to inhibit SARS-CoV-2 M^{pro} with high selectivity and with minimal off-target effects in humans^{20,29}.

The key role of M^{pro} for the replication and viability of SARS-CoV-2 has naturally led to the search for novel inhibitors of the protease. Perhaps the most promising of these are the peptidomimetic compounds developed by Pfizer, inspired by PF-00835231 (IC₅₀ = 4-8 nM against SARS-CoV-2 M^{pro}) that was originally developed against SARS-CoV-1 M^{pro}, which possesses high homology to the SARS-CoV-2 protease³⁰⁻³². Specifically, a phosphate prodrug of this inhibitor (PF07304814) has recently completed a phase 1b trial (clinical trials identifier NCT04535167). A second-

generation orally available peptidomimetic M^{pro} inhibitor developed by Pfizer (PF-07321332, Figure 1) has also recently entered phase 2 clinical trials for treatment of COVID-19 (clinical trials identifier NCT04960202)³³. Both molecules possess a γ -lactam mimic of the glutamine (Gln) residue found at the P₁ position in physiological cleavage sites and also inhibit SARS-CoV-2 M^{pro} via a covalent mechanism through electrophilic warheads embedded within the inhibitors³⁰. Specifically, PF-00835231 possesses a hydroxymethyl ketone moiety, while PF-07321332 contains a nitrile warhead, both of which react with the catalytic cysteine (Cys145) to inactivate the protease. In addition to these clinical candidates, a number of other peptidomimetic inhibitors of SARS-CoV-2 M^{pro} are currently under preclinical investigation³⁴⁻³⁷, including repurposed drugs such as boceprevir, a serine protease inhibitor approved in 2011 for the treatment of hepatitis C^{35,38,39}, and the feline anticoronaviral drug GC376^{35,38}.

Macrocyclic peptides are attractive chemotypes for medicinal chemistry efforts due to their ability to bind targets with high affinity and selectivity, whilst exhibiting greater proteolytic stability and membrane permeability than their linear counterparts⁴⁰⁻⁴². In this work we describe several potent cyclic peptide inhibitors of the SARS-CoV-2 M^{pro}, identified through the use of the Random nonstandard Peptide Integrated Discovery (RaPID) technology, which couples mRNA display with flexizyme-mediated genetic code reprogramming⁴²⁻⁴⁴. Importantly, we also report a crystal structure of the SARS-CoV-2 M^{pro} dimer bound to our most potent cyclic peptide inhibitor that highlights the residues important for binding both at the catalytic site and across the dimer interface. Finally, we demonstrate that three of the cyclic peptides identified exhibited antiviral

activity against SARS-CoV-2 *in vitro*, with an additional peptide gaining antiviral activity upon conjugation to a cell penetrating peptide.

Results

Selection against chemically cross-linked SARS-CoV-2 M^{pro} homodimer. In order to identify cyclic peptide inhibitors of SARS-CoV-2 M^{pro}, we sought to utilize RaPID, which allows the screening of $>10^{12}$ cyclic peptides for affinity against a protein target of interest immobilized on magnetic beads. However, functional SARS-CoV-2 M^{pro} is a homodimer with relatively weak affinity between the monomers ($K_D = 2.5 \mu M$), and we hypothesized that the protein may exist in a predominantly monomeric (i.e. inactive) form when immobilized on magnetic beads²⁷. Consistent with this, we found that the catalytic activity (measured by fluorescent substrate cleavage) of C-terminally His₆-tagged M^{pro} was significantly diminished (ca. 30%) when immobilized on His Pull-down DynabeadsTM compared to that of the wild-type protein in solution (**Supplementary Fig. 1a**); this indicated that a significant proportion of the immobilized M^{pro} was unable to form an active homodimer. We therefore investigated the use of a chemical cross-linking strategy to covalently lock SARS-CoV-2 M^{pro} in the homodimeric state to ensure that RaPID resulted in selection of ligands to the catalytically active form of the protease. The presence of proximal lysine residues near the dimer interface in the crystal structure of SARS-CoV-2 M^{pro} (PDB: 6Y2E) prompted us to investigate a number of amine-reactive bifunctional cross-linkers of differing lengths for this purpose²⁷. Efficient cross-linking was obtained by disuccinimidyl glutarate (DSG) (**Figure 2a and 2b**), with subsequent mapping by LC-MS/MS revealing intermolecular cross-links between lysine residues on the different monomer chains (K97-K97*, K97-K90*). We also observed cross-links between K12 and K97,

which are positioned in close proximity on different monomers of the dimer, as well as within the same monomer unit; however, in this case we could not differentiate between intermolecular or intramolecular cross-links by mass spectrometry (**Supplementary Fig. 1b**). Importantly, cross-linked M^{pro} exhibited catalytic activity comparable to that of wild-type M^{pro} (in solution), following immobilization on magnetic beads (**Supplementary Fig. 1a**). Based on these data, we moved on to RaPID selections against the cross-linked SARS-CoV-2 M^{pro} with a view to discovering novel cyclic peptide inhibitors.

To select for peptide binders of SARS-CoV-2 M^{pro} we performed parallel display using the RaPID mRNA display technology, employing peptide libraries initiated with either *N*-chloroacetyl-L-tyrosine or *N*-chloroacetyl-D-tyrosine to induce thioether macrocyclization through reaction with the side chain of a downstream cysteine residue. Tyr was chosen as the initiating amino acid due to its high translation efficiency in place of *N*-formyl-Met^{45,46}, while the use of D-Tyr in addition to L-Tyr expanded the accessible chemical space of the library. For each selection, a semi-randomized DNA library of >10¹³ unique sequences were transcribed into mRNA followed by ligation to a puromycin linker (**Figure 2c**). Translation of the mRNA-puromycin hybrids *in vitro*, followed by reverse transcription, yielded a library of peptide-mRNA:cDNA conjugates that were counter-selected against His Pull-down DynabeadsTM (to remove bead-binding peptides) before an affinity selection step, in which the libraries were panned against bead-immobilized cross-linked SARS-CoV-2 M^{pro}. PCR of the bead-bound fraction yielded an enriched DNA library that was used as the starting point for the subsequent rounds of selection (**Figure 2c**). Following nine iterative rounds of this process, next-generation sequencing of the final DNA library

identified several peptide sequences predicted to bind to SARS-CoV-2 M^{pro} based on high relative abundance (**Supplementary Fig. 2**).

Fig. 2. a Scheme for cross-linking of SARS-CoV-2 M^{pro} with disuccinimidyl glutarate (DSG). **b** SDS-PAGE gel of cross-linking reaction of M^{pro}. Lane 1: Reaction of SARS-CoV-2 M^{pro} (25 μM) with DSG (10 equivalents relative to M^{pro} monomer) for 1 h at 37°C in aqueous buffer (20 mM HEPES, 100 mM NaCl, pH 7.6). Lane 2: SARS-CoV-2 M^{pro}. Lane 3: Mark12TM ladder. **c** Workflow for the discovery of macrocyclic peptide ligands of cross-linked SARS-CoV-2 M^{pro} using RaPID mRNA display.

Synthesis and *in vitro* evaluation of cyclic peptide inhibitors of SARS-CoV-2 M^{pro}.

We selected eight peptides from the enriched libraries for synthesis (five L-Tyr initiated

and three D-Tyr initiated) based on their abundance in the final DNA library and diversity in sequence (**Figure 3a** and **3b**). Peptides **1–8** were synthesized *via* Fmoc-strategy solid-phase peptide synthesis on Rink amide resin with the N-terminal L- or D-Tyr residue derivatized with chloroacetic acid to facilitate thioether cyclization with the thiol of a downstream cysteine residue. Deprotection and cleavage from resin followed by cyclization provided the target thioether-linked cyclic peptides **1–8** following purification by reverse-phase HPLC. The synthetic cyclic peptides were next evaluated for inhibitory activity against SARS-CoV-2 M^{pro} using a previously reported fluorescence resonance energy transfer (FRET)-based assay^{27,47}. Gratifyingly, we observed potent inhibition of proteolytic activity for six of the eight peptides in the series, with IC₅₀ values for peptides **1–6** ranging from 0.070 – 12.7 μM (**Figure 3c**). Interestingly, despite being enriched in the selection, lariat peptide **7** and head to tail cyclic peptide **8** did not show appreciable inhibition of the protease (IC₅₀ > 50 μM). The most potent SARS-CoV-2 M^{pro} inhibitor in the series was cyclic peptide **1**, which exhibited an IC₅₀ of 70 ± 18 nM and a K_i of 14 ± 3 nM against the protease. Given the superior inhibitory potency of this molecule over the other cyclic peptides that emerged from the RaPID screen (>15-fold), this molecule was selected for further structure-activity investigations.

Fig. 3. a Synthesis of cyclic peptides **1–8** via Fmoc-SPPS (see Supplementary methods for full synthetic details). **b** Sequences of peptides **1–8** in one letter amino acid code, thioether cyclization is represented as a black line. **c** *In vitro* inhibition data of SARS-CoV-2 M^{pro} for peptides **1–6** with associated IC₅₀ and K_i values ±SEM. NB: cyclic peptides **7** and **8** showed no inhibition of SARS-CoV-2 M^{pro} at 50 μM, ND = not determined.

Probing the binding interaction of lead cyclic peptide 1

We next performed experiments to probe the key interactions of lead cyclic peptide inhibitor **1** with SARS-CoV-2 M^{pro}. Given the critical role dimerization plays in M^{pro} catalytic activity, we first investigated the effect of **1** on the monomer-dimer equilibrium of M^{pro} by multi-angle laser light scattering with size exclusion chromatography (SEC-MALLS). Interestingly, the addition of two molar equivalents of peptide **1** to a solution of wild-type M^{pro} (present in an approximately 1:1 ratio of monomer:dimer) resulted in a shift in the dimer-monomer equilibrium to afford a solution of exclusively homodimeric protease (**Figure 4a**). This data indicates that cyclic peptide **1** stabilizes the homodimeric form of SARS-CoV-2 M^{pro} and suggests that peptide **1** may bind to the active site present in the dimeric, but not monomeric forms of the protein. To corroborate this data, we used 3D NMR spectroscopy (specifically TROSY-HNCO spectra) to analyze SARS-CoV-2 M^{pro} after titration with peptide **1** (**Figure 4b**). This revealed shifts of backbone NMR resonances of residues near the active site, consistent with binding of peptide **1** at this location. However, shifts were also observed for residues located far from the active site (**Supplementary Fig. 3**), suggesting that the protein responds to the binding of **1** with global allosteric changes. In contrast, titration of the obligate monomeric mutant R298A of SARS-CoV-2 M^{pro} with peptide **1** led to no shifts in NMR resonances (**Supplementary Fig. 3**, see **Supplementary Tables 1 and 2** and **Supplementary Fig. 4** for full assignment of monomeric M^{pro} mutant). Notably, R298, which is mutated in the obligate monomer, is buried within the dimer interface in the structure of the wild-type protease and therefore the lack of binding of **1** to the monomeric protein is unlikely to be owing to a loss of interactions with this amino acid residue. Taken together, these SEC-MALLS and

NMR data therefore suggest that peptide 1 selectively binds to and inhibits the active homodimer of SARS-CoV-2 M^{pro} but does not bind to the inactive monomer.

Fig. 4. a SEC-MALLS of SARS-CoV-2 M^{pro} with and without peptide 1. SARS-CoV-2 M^{pro} (red line) exists in equilibrium between monomeric and homodimeric forms giving rise to two peaks (in ca. 1:1.25 ratio of monomer to dimer) in the size-exclusion chromatogram at a concentration of 1 μM in aqueous buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl). MALLS analysis indicates an approximate molecular weight of 36 kDa

for the monomer (calculated molecular weight = 33.8 kDa). After addition of two molar equivalents of **1**, M^{pro} converges predominantly to a homodimer (blue line) with a MALLS reading of 66 kDa (calculated MW of SARS-CoV-2 M^{pro} dimer = 67.6 kDa, MW of peptide **1** = 1874 Da) indicative of formation of the SARS-CoV-2 M^{pro} homodimer upon binding **1**. **b Cyclic peptide inhibitor 1 binds to the dimeric form of SARS-CoV-2 M^{pro}**. Overlay of projections onto the ¹⁵N-¹H plane of 3D TROSY-HNCO spectra of 0.3 mM solutions of ¹⁵N/¹³C/²H-labelled wild-type M^{pro}. Blue and red contour lines show the spectra recorded in the absence and presence of equimolar inhibitor **1**, respectively. Assignments are shown for peaks that shift or disappear in response to the inhibitor. **c Monitoring of the cleavage of cyclic peptide inhibitor 1 with SARS-CoV-2 M^{pro}**. MALDI-TOF mass spectrum of cyclic peptide **1** (top spectrum). Negligible cleavage of **1** was observed following incubation with SARS-CoV-2 M^{pro} under standard assay conditions (25 nM SARS-CoV-2 M^{pro}, 5 μM **1**, 20 mM Tris-HCl pH 7.6, 100 mM NaCl, 1 mM DTT, 1 mM EDTA) for 1 h at 37 °C (middle spectrum). Slow cleavage of **1** was observed (ca. ~30% after 1 h) in the presence of a high concentration of SARS-CoV-2 M^{pro} to 2.5 μM (bottom spectrum). **d Inhibitory activity of alanine mutants of lead cyclic peptide 1**. Sequences and associated inhibitory constants for peptide **1** analogues, whereby all polar residues within the randomized region of **1** were each systematically mutated to alanine and their inhibitory activity against SARS-CoV-2 M^{pro} assessed (Cys14 was not mutated as this residue is required for cyclization, Ala substitutions are shown in bold).

Cyclic peptide **1** includes the cleavage recognition motif of M^{pro} found in the natural viral protein substrates, namely a Gln in the P₁ position and a Leu in the P₂ position²⁷. It was therefore hypothesized that this Leu-Gln motif embedded within **1** mimics the

natural substrate and binds to the catalytic site of the protease. Importantly, this proposal is supported by the purely competitive inhibition mode observed for **1** (**Supplementary Fig. 5**). However, this also raises the possibility that the protease may be able to cleave **1** next to the Gln-Leu recognition sequence. To test this, we incubated **1** with wild type SARS-CoV-2 M^{pro} and assessed cleavage by MALDI-TOF-MS (**Figure 4c**). Peptide **1** exhibited notable resistance to proteolysis by M^{pro} with negligible cleavage of peptide **1** observed under standard assay conditions (25 nM M^{pro}) after 1 h. However, incubation of **1** with a high concentration of M^{pro} (2.5 μ M) resulted in slow cleavage of peptide **1**, with 30% peptide cleavage observed after 1 h. Analysis of the cleavage reaction by LC-MS/MS confirmed that proteolysis had indeed occurred between Gln3 and Tyr4 (**Supplementary Fig. 6a and 6b**). We also synthesized an authentic standard of the resulting cleavage product, which was verified to be identical to M^{pro}-cleaved **1** by LC-MS/MS (**Supplementary Fig. 6b**). Finally, we assessed whether the linear peptide product resulting from M^{pro} cleavage of **1** possessed inhibitory activity against SARS-CoV-2 M^{pro}. Interestingly, the peptide exhibited an IC₅₀ of 23.2 ± 5 μ M, ~330-fold higher than the IC₅₀ of 70 nM for **1** (**Supplementary Fig. 6c**). This two-orders of magnitude loss in activity upon linearization suggests that the conformation of cyclic peptide **1** is pre-organised for optimal interaction with the protease.

In order to assess the importance of each residue in **1** for M^{pro} inhibitory activity, we systematically replaced all residues in the peptide with alanine (except alkyl side chain- containing amino acids Ala5, Val6 and Leu7) and determined the inhibitory activity of the resulting mutants against M^{pro} (**Figure 4d, Supplementary Fig. 7**). Consistent with the known recognition sequence for SARS-CoV-2 M^{pro} and supported

by the mass spectrometry results described above, mutation of either Leu2 or Gln3 to Ala (that would be predicted to bind in the S2 and S1 recognition sites, respectively) led to more than two orders of magnitude reduction in inhibitory activity. Remarkably, mutation of Tyr4 (which would be predicted at P₁') also led to a significant loss in inhibitor activity (IC₅₀ = 1.9 μM); while this is consistent with the established importance of aromaticity in P₁', small residues such as alanine are often found in substrates at this position, and the dramatic reduction in inhibitory potency was therefore unexpected^{27,28}. Interestingly, mutation of Arg11, which is distal from the most prominent recognition residues of the cyclic peptide, also led to a significant reduction in inhibitory activity (IC₅₀ = 3.4 μM) suggesting that this residue makes important interactions with the protease and/or serves a crucial role in the adoption of the active conformation of the cyclic peptide. In contrast, mutation of Arg8, His9, Lys10, Arg12 or Glu13 resulted in equipotent activity or only a modest reduction in inhibitory activity (IC₅₀ values = 90–390 nM).

Co-crystallization of peptide 1 and a selenoether analogue with SARS-CoV-2 M^{pro}. In order to further interrogate the binding mode of inhibitor **1**, we used X-ray diffraction to solve the co-crystal structure of the SARS-CoV-2 M^{pro}-**1** complex to 3.4 Å. The limited resolution of the structure hindered the interpretation of the electron density within the active site. We therefore also solved the structure of a synthetic selenoether analogue of **1** (**Se-1**) in complex with SARS-CoV-2 M^{pro}, in which the cysteine residue involved in thioether peptide macrocyclization was replaced with selenocysteine⁴⁹; we rationalized that the greater electron density of Se in this analogue would aid its placement in the active site. Importantly, **Se-1** displayed identical inhibitory activity as **1** (K_i of **Se-1** = 17 nM vs 14 nM for **1**), suggesting its

interactions with the protein are also very similar to parent inhibitor **1** (**Supplementary Fig. 8**). A crystal of this SARS-CoV-2 M^{pro}-**Se-1** complex diffracted to 2.19 Å resolution in the same space group and with similar cell dimensions as the SARS-CoV-2 M^{pro}-**1** complex, thus providing a detailed view of the interactions between the inhibitor and the protease (**Supplementary Table 3**). The crystal dimensions and packing are unique and not observed in any of the hundreds of deposited SARS-CoV-2 M^{pro} structures, with a notably long *b* axis. There were four protein subunits within the asymmetric unit comprising two physiological dimers (**Figure 5a**).

Although the peptide is not fully resolved in the structure, 9-13 residues were observable in various chains with good electron density; the selenoether linkage and the first 5 residues (D-Tyr1, Leu2, Gln3, Tyr4, Ala5) were very stable, with the remainder of the peptide (7-12) appearing somewhat disordered (**Figure 5b**). The peptide displayed a consensus binding pose across three subunits of the crystal structure. The fourth subunit displayed an alternative binding pose in which Leu2 and Gln3 were present in identical positions, but the flanking D-Tyr1 and Tyr4 residues adopted alternative conformations (**Supplementary Fig. 9a and 9b**). Closer inspection revealed a non-Pro *cis*-peptide bond between Tyr4 and Ala3 in chains A, C, D, which results in a tight kink in the helix, while in chain B this bond remains in the *trans*-configuration (**Supplementary Fig. 9c and 9d**). The Gln3 of the peptide occupies the canonical S₁ subsite of the protease (following the numbering by Lee *et al.*)⁵⁰ in every peptide:protein complex, facilitated by interactions with His163, Glu166, and Asn142 (**Figure 5b**). Likewise, Leu2 is always bound in the S₂ subsite comprised of His41, His164 and Gln189. D-Tyr1 is in the solvent-exposed S₃ site bordered by Gln189, Ala191 and Pro16, while the peptide twists at the selenoether/D-Tyr1 linkage,

turning away from the canonical S₄ site to be positioned adjacent to Glu166. Tyr4 is therefore positioned in the S₁' (P₁') position in ¾ chains, fully consistent with the slow proteolysis between Gln3 and Tyr4 observed by mass spectrometry (**Figure 4c**). In chain B, the peptide backbone and Tyr4 have swapped positions (**Supplementary Fig. 9a and 9b**).

To obtain a better understanding of the interaction of the full peptide with the protein in the absence of crystal contacts, we modelled the missing residues as accurately as possible with the available density and used this structure as the starting point for triplicate 333 ns molecular dynamics (MD) simulations⁵¹ (~1 µs total simulation time; **Figure 5c, Supplementary Fig. 10**). The results were consistent with the crystal structure, i.e. residues 1-4 were all relatively stable within their respective binding pockets within the substrate cleft of SARS-CoV-2 M^{pro}, with the mobility of the peptide increasing on either side of these residues (**Figure 5d**). Notably, the simulations showed regular transient interactions between the peptide and the second chain of the physiological dimer.

The structure and MD simulations also provide a molecular explanation for the inhibitory activity of the alanine mutants. Specifically, the structure shows that the central interactions are formed by Leu2 and Gln3, consistent with the large reductions in activity when these positions are mutated to Ala, while D-Tyr1 and Tyr4 also form significant interactions with the protease on either side of these residues. Other positions (8, 9, 10, 12 and 13) that were observed to have little influence on inhibitory activity are either disordered or solvent-exposed. MD simulations suggest that Arg11, the mutation of which to Ala had a significant effect on inhibition (**Figure 4d**), makes

contacts across the dimer interface (**Supplementary Fig. 9e**), interacting with the side chain and main chain carbonyl of Q256*. This is consistent with electron density showing Arg11 within hydrogen bonding distance of the neighbouring chain B in one dimer (**Supplementary Fig. 9f**). These observations are consistent with the SEC-MALLS and NMR data that shows the peptide binds the dimeric form of SARS-CoV-2 M^{pro} exclusively for its mode of inhibition (**Figure 4a and 4b**).

Fig. 5. Structural analysis of the SARS-CoV-2 M^{pro}-Se-1 complex (PDB ID: 7RNW). a The asymmetric unit of the SARS-CoV-2 M^{pro}-Se-1 crystal structure, containing two physiological dimers shown as blue and green. The Se-1 peptide is

shown in a sphere representation. **b** Zoom of the active site with electron density (Polder omit map, contoured to 3.5σ) showing bound peptide (residues 7–12 disordered). Key active site residues of SARS-CoV-2 M^{pro} are shown as sticks, and key positions of the **Se-1** peptide (residues 1-4) are labelled in bold text. **c** Representative conformations (300 ns) from each of the triplicate simulations suggest that Tyr1, Leu2, Gln3 and Tyr4 are stably bound in the active site of SARS-CoV-2 M^{pro}, while the remainder of the peptide is more mobile and makes transient interactions across the dimer interface. **d** **Se-1** peptide bound to chains A, C and D coloured by B-factor, showing the residues 1-4 are stable and bound within the S₃, S₂, S₁ and S₋₁ subsites, respectively, whereas the polar half of the peptide (Arg8-Glu13) is either too disordered to accurately model, or is modelled with high B-factors.

Antiviral activity of cyclic peptide SARS-CoV-2 M^{pro} inhibitors. Given the promising inhibitory activity of cyclic peptides **1–6** against SARS-CoV-2 M^{pro}, we next assessed the antiviral activity of the four most potent peptides against SARS-CoV-2 *in vitro*. Specifically, we used ACE-2 and TMPRSS2 overexpressing HEK293T cells infected with SARS-CoV-2 to assess each peptide. The latter cell line is hyper-permissive to SARS-CoV-2 infection with extensive viral syncytia forming after 18 hours post viral infection. Dose-dependent sigmoidal inhibition curves can be generated through enumeration of remaining single cell nuclei using a live nuclear dye and a standard high content fluorescent microscope. In this setting, we observed a dose-dependent protection for three of the four peptides, with EC₅₀ values ranging from 11.8 – 33.1 μ M. However, peptide **1**, the most active compound in the enzyme inhibition assay, was inactive up to a concentration of 50 μ M, most likely due to low cell permeability owing to the large number of polar and charged residues in this

439 molecule. To facilitate cell entry of **1**, we therefore conjugated its C-terminus to
 440 penetratin, a 16 amino acid cell-penetrating peptide (CPP) derived from the *Drosophila*
 441 Antennapedia homeodomain⁵² (to afford **pen-1**). It should be noted that during the
 442 selection process the C-terminus of the peptides is conjugated to a large mRNA:cDNA
 443 tag, and therefore addition of a CPP to the C-terminus of the peptide (to afford **pen-1**)
 444 was deemed unlikely to affect binding and inhibition of M^{pro}. Pleasingly, **pen-1** had
 445 equivalent *in vitro* inhibitory activity against SARS-CoV-2 M^{pro} to peptide **1** (K_i = 9 nM
 446 for **pen-1** vs 14 nM for **1**, **Supplementary Fig. 11**) indicating that addition of a CPP to
 447 the C-terminus of the peptide did not affect binding and inhibition of the protease. We
 448 also prepared a penetratin conjugate of cyclic peptide **6**, the second most active
 449 peptide in the enzyme inhibition assay (K_i = 360 nM). While parent cyclic peptide **6**
 450 showed antiviral activity in the cell-based assay (EC₅₀ of 33.1 μ M), we anticipated that
 451 the conjugation to penetratin might further improve potency. We also assessed the
 452 cytotoxicity of **1**, **pen-1**, **6** and **pen-6** on HEK293-ACE2-TMPRSS2 cells, with none of
 453 the compounds affecting cell viability at a concentration of 50 μ M (**Supplementary**
 454 **Fig. 12**). Pleasingly, we found that **pen-1** and **pen-6** both exhibited significantly
 455 improved antiviral activity, with EC₅₀ values of 15.2 μ M and 6.6 μ M, respectively
 456 (**Figure 6**). As expected, the marked improvement in antiviral activity of peptide **1** by
 457 conjugation to the penetratin CPP correlated with enhanced cellular uptake, whereby
 458 LC-MS/MS analysis showed a 5.5-fold increase in levels of **pen-1** compared to **1** in
 459 cell lysates (**Supplementary Fig. 13**).

Fig. 6. Antiviral activity of peptides 1, 2, 5, 6 and penetratin conjugates of 1 and 6 (pen-1 and pen-6). HEK293-ACE2-TMPRSS2 cells were incubated with varying concentrations of cyclic peptide M^{pro} inhibitors and infected with SARS-CoV-2. Inhibition curves and 50% effective concentrations (EC₅₀) were determined by non-linear regression analysis using GraphPad Prism. Data are the means ± SD of experiments performed in quadruplicate.

Discussion

In summary, we discovered macrocyclic peptide inhibitors of SARS-CoV-2 M^{pro} using RaPID mRNA display technology. A protein cross-linking strategy was employed to generate a covalently locked catalytically active dimer, facilitating display selection against solid-supported SARS-CoV-2 M^{pro}. This enabled the discovery of potent inhibitors of M^{pro}, including the 14-residue cyclic peptide **1** with a K_i of 14 nM that represents one of the most potent inhibitors reported against the protease to date. SEC-MALLS and NMR spectroscopic studies revealed that this molecule inhibits SARS-CoV-2 M^{pro} by exclusively binding to the catalytic protease dimer, highlighting the importance of controlling the nature of the immobilized protein construct for display

selections; in this case by covalently cross-linking the M^{pro} homodimer. Co-crystal structures of M^{pro} and **1** and the selenocysteine derivative **Sec-1** solved by X-ray crystallography revealed the canonical interaction between a Gln residue and subsite S₁, flanked by hydrophobic residues. This half of the peptide was relatively stable, while the hydrophilic face (8–12) was relatively mobile, interacting with solvent and across the dimer interface to facilitate an exclusive dimer binding mode for the inhibitor. The strain imposed by cyclization prevents the peptide from adopting a fully relaxed conformation and likely explains the very slow turnover, despite the inhibitor possessing canonical residues that bind to the key recognition subsites of the protease. Thus, cyclization of the peptide has essentially converted a peptide substrate into a highly potent SARS-CoV-2 M^{pro} inhibitor, highlighting a key benefit of the cyclic peptides that emerge from RaPID screens in medicinal chemistry studies. Several of the cyclic peptides also exhibited antiviral activity against SARS-CoV-2 *in vitro* and enhancing cellular uptake of cyclic peptides by conjugation to the CPP penetratin led to dramatic improvements in antiviral activity. This was especially pronounced for lead molecule **1**, which was ineffective without the covalent CPP tag but exhibited an EC₅₀ of 15.9 μM following fusion to penetratin. Taken together, these compounds now serve as bona fide starting points for the rational design of peptide- or peptidomimetic-based antivirals for COVID-19 that target SARS-CoV-2 M^{pro}. Moreover, the crosslinked protease selection strategy developed as part of this work is expected to inform medicinal chemistry campaigns targeting other dimeric coronaviral main proteases. Future work in our laboratories will involve lead optimization of the discovered peptides for improved activity and cell permeability both by rational design and high-throughput affinity maturation^{53,54} with a view to developing molecules with potent antiviral activity *in vivo*.

Methods

Plasmid construction

The gene of the SARS-CoV-2 main protease (M^{pro})²⁷ was cloned in between the *Nde*I and *Xho*I sites of the T7 vector pET-47b (+). The construct contains the M^{pro} self-cleavage-site (SAVLQ↓SGFRK; arrow indicating the cleavage site) at the N-terminus. At the C-terminus, the construct contains a modified PreScission cleavage site (SGVTFFQ↓GP) connected to a His₆-tag. All plasmid constructions and mutagenesis were conducted with cloning and QuikChange protocols using mutant T4 DNA polymerase.

Protein expression

Wild-type M^{pro} was expressed in BL21 DE3 cells transformed with the desired plasmid. Protein expression was conducted in a Labfors 5 bioreactor (INFORS HT, Switzerland). After induction, the culture was grown at 18 °C overnight for protein expression. Cells were harvested by centrifugation at 5,000 g for 15 minutes and lysed by passing twice through a Emulsiflex-C5 homogenizer (Avestin, Canada). The lysate was centrifuged at 13,000 g for 60 minutes and the filtered supernatant was loaded onto a 5 mL Ni-NTA column (GE Healthcare, USA) equilibrated with binding buffer (50 mM Tris-HCl pH 7.5, 300 mM NaCl, 5 % glycerol). The protein was eluted with elution buffer (binding buffer containing, in addition, 300 mM imidazole) and the fractions were analyzed by 12% SDS-PAGE. PreScission cleavage and TEV cleavage were conducted in binding buffer in the presence of 1 mM DTT with a protein-to-protease ratio of 100:1. Following cleavage of the His₆-tag, the buffer was exchanged to 20 mM HEPES-KOH pH 7.0, 150 mM NaCl, 1 mM DTT, 1 mM EDTA. All samples were analyzed by mass

spectrometry using an Orbitrap Fusion Tribrid mass spectrometer (Thermo Scientific, USA) coupled with an UltiMate S4 3000 UHPLC (Thermo Scientific, USA).

Protein cross-linking

C-terminally His-tagged SARS-CoV-2 M^{pro} (25 µM) was incubated with disuccinimidyl glutarate (H6/D6, Creative Molecules Inc.) (250 µM) in aqueous buffer (20 mM HEPES pH 7.6, 100 mM NaCl) at 37 °C for 1 h before the reaction was quenched by dilution of the protein to 10 µM (monomer concentration) by addition of aqueous buffer containing 20 mM Tris-HCl pH 7.6, 100 mM NaCl. Cross-linking efficiency was analyzed by 12% SDS-PAGE and stained with SyproTM Ruby.

Proteomics to determine crosslink sites on SARS-CoV-2 M^{pro}

Crosslinked proteins were digested with trypsin and peptides desalted as described previously⁵⁵. Peptide mixtures were analyzed by LC-MS/MS using data dependent acquisition with HCD fragmentation on a Thermo Orbitrap Eclipse mass spectrometer. Crosslinked peptides were identified with a 2% false-discovery rate using the Byonic search engine (Protein Metrics) and a custom database containing the 3CL protein sequence and common proteomics contaminants (e.g. trypsin, albumin, keratins, etc.). DSG crosslinks and hydrolysis products were allowed for the 3CL protein only. Carbamidomethylation was specified as a fixed modification on C, while oxidation of methionine, deamidation of N/Q and pyro-Glu for N-terminal Q/E were variable modifications. Plots of crosslinked residues were generated using UCSF Chimera 1.15rc with the 3CL dimer structure (PDB: 6Y2E).

RaPID mRNA display selection

Library preparation and display selection were performed as described previously with slight modifications^{44,56-58}. Briefly, an mRNA library comprising an AUG start codon, 4–15 NNS (N = A, C, G or T; S = C or G) codons, a TGC cysteine codon and a 3' region encoding a linking sequence (Gly-Asn-Leu-Ile) with mRNA lengths pooled proportional to theoretical diversity was prepared as described previously. The mRNA library was then ligated to a puromycin-linked oligonucleotide using T4 ligase. The puromycin-linked mRNA was then translated using the PURExpress Δ RF kit (New England Biolabs) with the addition of release factors 2 and 3 per the manufacturer's instructions along with a "Solution A" containing 19 amino acids (-Met) (prepared as previously described) and either tRNA^{fMet}_{Ini}-N-chloroacetyl-L-Tyr or tRNA^{fMet}_{Ini}-N-chloroacetyl-D-Tyr (25 μ M) to initiate translation and facilitate peptide macrocyclization. Translation was performed at 100 μ L scale for the first round and 2.5 μ L scale for subsequent rounds. Following translation, ribosomal denaturation was performed by addition of EDTA before reverse transcription with RNase H- reverse transcriptase. After exchange into selection buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 0.1% Tween-20) through G-25 resin the peptide-mRNA:cDNA libraries were incubated with cross-linked linked SARS-CoV-2 M^{pro} (200 nM) immobilized on DynabeadsTM His-tag Isolation & Pulldown (Life Technologies) at 4 °C for 30 min after which the beads were washed with ice-cold selection buffer (3 x 100 μ L). The beads were then suspended in aqueous 0.1 vol.% TritonTM X-100 (100 μ L) and heated to 95 °C for 5 min to recover binding peptides. The recovered cDNA was amplified by PCR, purified by ethanol precipitation and transcribed with T7 RNA polymerase to yield an enriched mRNA library to enter the following round of selection. Subsequent rounds were performed in the same manner with the addition of six negative selection steps against beads alone to remove bead binding sequences. Sequencing of recovered cDNA from rounds

seven through nine was conducted using an iSeq high-throughput sequencer (Illumina).

Cyclic peptide synthesis

For materials and methods, full synthetic details and characterization (HPLC, ESI-MS) of all peptides, see Supplementary Methods and Supplementary Figures 14-19.

SARS-CoV-2 M^{pro} activity assay

The M^{pro} inhibition assay protocol was adapted and modified from^{27,47} and was carried out with a reaction volume of 100 μ L in 96-well-plates (black, polypropylene, U-bottom; Greiner Bio-One, Austria) using an aqueous buffer composed of 20 mM Tris-HCl pH 7.3, 100 mM NaCl, 1 mM EDTA, 1 mM DTT. All measurements were performed in triplicate. The compounds were incubated with recombinant SARS-CoV-2 M^{pro} for 10 min at 37 °C. The enzymatic reaction was initiated by addition of the FRET substrate (DABCYL)-KTSAVLQ↓SGFRKM-E(EDANS)-NH₂ (Mimotopes, Australia). The final concentrations for IC₅₀ determination amounted to 25 nM M^{pro} and 25 μ M substrate, with inhibitor concentrations ranging from nanomolar to micromolar. The final concentrations for K_i determination amounted to 12.5 nM enzyme and 10, 20, 35 and 50 μ M substrate, with a control and three inhibitor concentrations ranging from nanomolar to micromolar. The fluorescence signal was monitored by a fluorophotometer (Infinite 200 PRO M Plex; Tecan, Switzerland) for 5 min at 490 nm, using an excitation wavelength of 340 nm. Initial velocities were derived from the linear range of the enzymatic reaction. For IC₅₀ determination, 100% enzymatic activity was defined as the initial velocity of control triplicates containing no inhibitor and the percentage of inhibition was calculated in relation to 100% enzymatic activity. An

EDANS standard curve generated as described by Ma *et al.*³⁵ was used to convert fluorescence intensity to the amount of cleaved substrate (calibration curve). Data sets were analyzed with Prism 9.2 (Graph Pad Software, USA) to generate dose-response curves and calculate IC₅₀ values, as well as to generate Michaelis-Menten curves and calculate K_i values assuming the appropriate inhibition mode.

Crystallography

To obtain a crystallographic complex of SARS-CoV-2 M^{pro} and **1**, the peptide was dissolved in DMSO to a stock concentration of 50 mM. Protein, prepared as described above, was diluted to 4 mg/mL in TBS (50 mM Tris-HCl pH 7.5, 300 mM NaCl) and incubated with 2.5-fold molar excess (~300 μ M) **1** for 2 hours on ice to saturate the binding sites of the protease. The complex was briefly buffer exchanged to remove unbound peptide using an Amicon Ultra 0.5 mL centrifugal filter with a 10 kDa MWCO to a final protein concentration of 4 mg/mL. This complex was used for high-throughput crystallography trials at 18 °C using drop sizes of 0.5 μ L protein and 0.5 μ L reservoir, which yielded a hit in the Index sparse matrix screen (Hampton Research). This hit was optimized but did not yield crystals of suitable diffraction quality; however, the crystals were able to be crushed and used to seed crystals with better morphology. Serial seeding was performed using a protein concentration of 1 mg/mL for several rounds until crystals of suitable diffracting quality were obtained, which were formed at 18 °C in a drop size of 1 μ L reservoir, 1.5 μ L protein, and 0.5 μ L seed stock against a reservoir solution of 25% PEG 3350, 0.1 M Bis-Tris pH 6.5, 0.3 M NaCl. Crystals were flash frozen without cryoprotecting and diffraction data were collected at 100 K using the MX2 beamline at the Australian Synchrotron. As the diffraction was not

adequate to unambiguously model the inhibitor into the crystal structure, we solved the co-crystal structure of the SARS-CoV-2 Mpro-**Se-1** complex. The complex was prepared the same as with **1** but with overnight incubation at 4 °C. The complex was diluted to a protein concentration of 1 mg/mL and crystallized at 18 °C in a drop size of 1 µL reservoir, 1.5 µL protein and 0.5 µL seed stock against a reservoir solution of 22% PEG 3350, 0.1 M Bis-Tris pH 6.0, 0.3 M NaCl. Crystals were seeded using crystal seeds of the SARS-CoV-2 Mpro-**1** crystals. Crystals formed as thin plates after ~4–5 days and were flash frozen without cryoprotection.

Diffraction data were collected at 100 K using the MX2 beamline at the Australian Synchrotron⁵⁹. Reflections collected were indexed and integrated using XDS⁶⁰ and scaled in Aimless (CCP4)⁶¹. The phase problem was overcome by molecular replacement in Phaser MR (CCP4)⁶², using PDB ID 7JKV as the search model. The structure was refined by iterative rounds of rebuilding in Coot⁶³, and twin refinement in Refmac^{64,65}. The crystal was twinned P21 (-h, -k, (k+l)); initially autoindexing as C221). Data collection and refinement statistics are given in **Supplementary Table 3**. The final structure was deposited in the Protein Data Bank (PDB: 7RNW).

SARS-CoV-2 Infectivity Assays

HEK293-ACE2-TMPRSS2 cells stably expressing human ACE2 and TMPRSS2 were generated as previously described⁶⁶. A high content fluorescence microscopy approach was used to assess the ability of the cyclic peptide M^{pro} inhibitors to protect cells from SARS-CoV-2 induced cytopathic effects in permissive cells. In brief, the engineered cell line succumbs to viral cytopathic effects after 6 to 18 hours post infection. Cytopathic effects can be enumerated, as cells and their nuclei collapse into

large syncytia after 18 hours of viral culture. The remaining cells outside of the syncytia increase in a dose dependent manner the lower the viral titers are and/or if a viral inhibitor is introduced within the culture.

For testing the cyclic peptides, compounds were initially diluted in cell culture medium (DMEM-5% FCS) to make 4x working stock solutions and then serially diluted further in the above media to achieve a 2-fold dilution series. On the day of the assay, HEK293-ACE2-TMPRSS2 cells were trypsinized, stained with Nucblue in suspension and then seeded at 16,000 cells in a volume of 40 μ L of DMEM-5% FCS per well in a 384-well plate (Corning #CLS3985). Diluted compounds (20 μ L) were added to the cells and the plates containing cells and compounds incubated for 1 hour at 37 °C, 5% CO₂. 20 μ L of virus solution at 8×10^3 TCID₅₀/mL⁶⁷ was then added to the wells and plates were incubated for a further 24 hours at 37 °C, 5% CO₂. Stained cells were then imaged using the InCell 2500 (Cytiva) high throughput microscope, with a 10 $\times$ 0.45 NA CFI Plan Apo Lambda air objective. Acquired nuclei were counted using InCarta high-content image analysis software (Cytiva) to give a quantitative measure of CPE. Virus inhibition/neutralization was calculated as %N= (D-(1-Q))x100/D, where; “Q” is the value of nuclei in test well divided by the average number of nuclei in untreated uninfected controls, and “D”=1-Q for the wells infected with virus but untreated with inhibitors. Thus, the average nuclear counts for the infected and uninfected cell controls get defined as 0% and 100% neutralization respectively. To account for cell death due to drug toxicity, cells treated with a given compound alone and without virus were included in each assay. The % neutralization for each compound concentration in infected wells was normalized to % neutralization in wells with equivalent amount of compound but without the virus to yield the final neutralization values for each condition. Inhibition curves and 50% (EC₅₀) effective concentrations were determined

by non-linear regression analysis using GraphPad Prism software (version 9.1.2, GraphPad software, USA).

Cytotoxicity and targeted proteomics on HEK293-ACE2-TMPRSS2 cell line

HEK293-ACE2-TMPRSS2 cells were maintained in Dulbecco's Modified Eagle Medium (DMEM) with 4.5 g/L D-Glucose, L-Glutamine and 110 mg/L sodium pyruvate (Gibco), supplemented with 10% foetal calf serum (Hyclone). Cells were sub-cultured between 70-90% confluency and tested regularly to ensure free of mycoplasma contamination. Cytotoxicity of **1**, **pen-1**, **6** and **pen-6**, was determined by an Alamar Blue HS (Invitrogen) cell viability assay as per manufacturer's instructions. Briefly, 5×10^4 cells were seeded into wells of a 96-well flat bottom culture plate (Corning) and once adhered, compounds or vehicle control (DMSO) were added at varying concentrations and incubated for 24 hours (37 °C, 5% CO₂). Alamar Blue HS cell viability reagent (Invitrogen) was added and the cells further incubated for 2-3 hours. Relative fluorescent units (RFU) were determined per well at ex/em 560/590 nm (Tecan Infinite M1000 pro plate reader). Increasing RFU is proportional to cell viability.

For targeted proteomics, cells in 6-well plates at 70% confluency were treated with vehicle control (DMSO) or inhibitors at 10 µM. At various time points (performed in triplicate), the cells were washed 3 times with 2 mL DPBS (Gibco), scraped and lysed in 4% SDC buffer (4% sodium deoxycholate, 0.1 M Tris-HCl pH 8.0) then immediately heated to 95 °C for 10 mins before freezing at -30 °C prior to processing for targeted proteomics⁶⁸. Cell lysates were digested to peptides as described previously. Peptides were analyzed by LC-MS/MS using a data-independent acquisition method as described previously.⁶⁹ Extracted ion chromatograms for fragment ions derived from

tryptic peptides were plotted using Xcalibur Qual Browser (Thermo Scientific). The area under the curve for each peak was used for quantification.

Data availability

The X-Ray co-crystal structure of SARS-CoV-2 M^{pro}-**Se-1** has been deposited to the PDB (PDB ID: 7RNW). All other data supporting the findings reported in this study are available in the article and its Supplementary Information.

References

- 1 Zhou, P. *et al.* A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **579**, 270-273, (2020).
- 2 Dong, E., Du, H. & Gardner, L. An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect. Dis.* **20**, 533-534, (2020).
- 3 Wall, E. C. *et al.* Neutralising antibody activity against SARS-CoV-2 VOCs B.1.617.2 and B.1.351 by BNT162b2 vaccination. *Lancet* **397**, 2331-2333, (2021).
- 4 Wang, Y. *et al.* Remdesivir in adults with severe COVID-19: a randomised, double-blind, placebo-controlled, multicentre trial. *Lancet* **395**, 1569-1578, (2020).
- 5 Beigel, J. H. *et al.* Remdesivir for the treatment of COVID-19 — final report. *N. Engl. J. Med.* **383**, 1813-1826, (2020).
- 6 Dyer, O. COVID-19: Remdesivir has little or no impact on survival, WHO trial shows. *BMJ* **371**, m4057, (2020).
- 7 Cao, B. *et al.* A trial of lopinavir–ritonavir in adults hospitalized with severe COVID-19. *N. Engl. J. Med.*, (2020).
- 8 Horby, P. W. *et al.* Lopinavir–ritonavir in patients admitted to hospital with COVID-19 (RECOVERY): a randomised, controlled, open-label, platform trial. *Lancet* **396**, 1345-1352, (2020).
- 9 Schreiber, G. The role of type I interferons in the pathogenesis and treatment of COVID-19. *Front. Immunol.* **11**, 595739, (2020).
- 10 Monk, P. D. *et al.* Safety and efficacy of inhaled nebulised interferon beta-1a (SNG001) for treatment of SARS-CoV-2 infection: a randomised, double-blind, placebo-controlled, phase 2 trial. *Lancet Respir. Med.* **9**, 196-206, (2021).
- 11 Boulware, D. R. *et al.* A randomized trial of hydroxychloroquine as postexposure prophylaxis for COVID-19. *N. Engl. J. Med.* **383**, 517-525, (2020).
- 12 Skipper, C. P. *et al.* Hydroxychloroquine in nonhospitalized adults with early COVID-19: a randomized trial. *Ann. Intern. Med.* **173**, 623-631, (2020).
- 13 Cavalcanti, A. B. *et al.* Hydroxychloroquine with or without azithromycin in mild-to-moderate COVID-19. *N. Engl. J. Med.* **383**, 2041-2052, (2020).
- 14 Tummino, T. A. *et al.* Drug-induced phospholipidosis confounds drug repurposing for SARS-CoV-2. *Science* **373**, 541-547, (2021).
- 15 Group, R. C. Dexamethasone in hospitalized patients with COVID-19—preliminary report. *N. Engl. J. Med.* **384**, 693-704, (2020).

- 750 16 Keller, M. J. *et al.* Effect of systemic glucocorticoids on mortality or mechanical
751 ventilation in patients with COVID-19. *J. Hosp. Med.* **15**, 489-493, (2020).
- 752 17 Hoffmann, M. *et al.* SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2
753 and is blocked by a clinically proven protease inhibitor. *Cell* **181**, 271-280,
754 (2020).
- 755 18 Bayati, A., Kumar, R., Francis, V. & McPherson, P. S. SARS-CoV-2 infects cells
756 after viral entry via clathrin-mediated endocytosis. *J. Biol. Chem.* **296**, 100306,
757 (2021).
- 758 19 Shang, J. *et al.* Cell entry mechanisms of SARS-CoV-2. *Proc. Natl. Acad. Sci.*
759 *U.S.A.* **117**, 11727-11734, (2020).
- 760 20 V'kovski, P., Kratzel, A., Steiner, S., Stalder, H. & Thiel, V. Coronavirus biology
761 and replication: implications for SARS-CoV-2. *Nat. Rev. Microbiol.* **19**, 155-170,
762 (2021).
- 763 21 Shin, D. *et al.* Papain-like protease regulates SARS-CoV-2 viral spread and
764 innate immunity. *Nature* **587**, 657-662, (2020).
- 765 22 Wu, Y. *et al.* Main protease of SARS-CoV-2 serves as a bifunctional molecule
766 in restricting type I interferon antiviral signaling. *Signal Transduction Targeted*
767 *Ther.* **5**, 221, (2020).
- 768 23 Fung, S. Y., Siu, K. L., Lin, H., Yeung, M. L. & Jin, D. Y. SARS-CoV-2 main
769 protease suppresses type I interferon production by preventing nuclear
770 translocation of phosphorylated IRF3. *Int. J. Biol. Sci.* **17**, 1547-1554, (2021).
- 771 24 Blanco-Melo, D. *et al.* Imbalanced host response to SARS-CoV-2 drives
772 development of COVID-19. *Cell* **181**, 1036-1045, (2020).
- 773 25 Hadjadj, J. *et al.* Impaired type I interferon activity and inflammatory responses
774 in severe COVID-19 patients. *Science* **369**, 718-724, (2020).
- 775 26 Anand, K. *et al.* Structure of coronavirus main proteinase reveals combination
776 of a chymotrypsin fold with an extra α -helical domain. *EMBO J.* **21**, 3213-3224,
777 (2002).
- 778 27 Zhang, L. *et al.* Crystal structure of SARS-CoV-2 main protease provides a
779 basis for design of improved α -ketoamide inhibitors. *Science* **368**, 409-412,
780 (2020).
- 781 28 Ullrich, S. & Nitsche, C. The SARS-CoV-2 main protease as drug target. *Bioorg.*
782 *Med. Chem. Lett.* **30**, 127377, (2020).
- 783 29 Anand, K., Ziebuhr, J., Wadhwani, P., Mesters, J. R. & Hilgenfeld, R.
784 Coronavirus main proteinase (3CLpro) structure: basis for design of anti-SARS
785 drugs. *Science* **300**, 1763-1767, (2003).
- 786 30 Boras, B. *et al.* Discovery of a Novel Inhibitor of Coronavirus 3CL Protease for
787 the Potential Treatment of COVID-19. *bioRxiv*, 2020.2009.2012.293498, (2021).
- 788 31 Hoffman, R. L. *et al.* Discovery of ketone-based covalent inhibitors of
789 coronavirus 3CL proteases for the potential therapeutic treatment of COVID-19.
790 *J. Med. Chem.* **63**, 12725-12747, (2020).
- 791 32 Vankadara, S. *et al.* A head-to-head comparison of the inhibitory activities of 15
792 peptidomimetic SARS-CoV-2 3CLpro inhibitors. *Bioorg. Med. Chem. Lett.* **48**,
793 128263-128263, (2021).
- 794 33 Vandyck, K. & Deval, J. Considerations for the discovery and development of
795 3-chymotrypsin-like cysteine protease inhibitors targeting SARS-CoV-2
796 infection. *Curr. Opin. Virol.* **49**, 36-40, (2021).
- 797 34 Dai, W. *et al.* Structure-based design of antiviral drug candidates targeting the
798 SARS-CoV-2 main protease. *Science* **368**, 1331-1335, (2020).

799 35 Ma, C. *et al.* Boceprevir, GC-376, and calpain inhibitors II, XII inhibit SARS-
800 CoV-2 viral replication by targeting the viral main protease. *Cell Res.* **30**, 678-
801 692, (2020).

802 36 Yang, K. S. *et al.* A quick route to multiple highly potent SARS-CoV-2 main
803 protease inhibitors. *ChemMedChem* **16**, 942-948, (2021).

804 37 Kitamura, N. *et al.* Expedited approach toward the rational design of
805 noncovalent SARS-CoV-2 main protease inhibitors. *J. Med. Chem.*, DOI:
806 10.1021/acs.jmedchem.1c00509, (2021).

807 38 Fu, L. *et al.* Both Boceprevir and GC376 efficaciously inhibit SARS-CoV-2 by
808 targeting its main protease. *Nat. Commun.* **11**, 4417, (2020).

809 39 Ghahremanpour, M. M. *et al.* Identification of 14 known drugs as inhibitors of
810 the main protease of SARS-CoV-2. *ACS Med. Chem. Lett.* **11**, 2526-2533,
811 (2020).

812 40 Vinogradov, A. A., Yin, Y. & Suga, H. Macrocyclic peptides as drug candidates:
813 recent progress and remaining challenges. *J. Am. Chem. Soc.* **141**, 4167-4181,
814 (2019).

815 41 Zorzi, A., Deyle, K. & Heinis, C. Cyclic peptide therapeutics: past, present and
816 future. *Curr. Opin. Chem. Biol.* **38**, 24-29, (2017).

817 42 Passioura, T. & Suga, H. A RaPID way to discover nonstandard macrocyclic
818 peptide modulators of drug targets. *Chem. Commun.* **53**, 1931-1940, (2017).

819 43 Peacock, H. & Suga, H. Discovery of de novo macrocyclic peptides by
820 messenger RNA display. *Trends Pharmacol. Sci.* **42**, 385-397, (2021).

821 44 Goto, Y., Katoh, T. & Suga, H. Flexizymes for genetic code reprogramming.
822 *Nat. Protoc.* **6**, 779-790, (2011).

823 45 Goto, Y. *et al.* Reprogramming the translation initiation for the synthesis of
824 physiologically stable cyclic peptides. *ACS Chem. Biol.* **3**, 120-129, (2008).

825 46 Fujino, T., Goto, Y., Suga, H. & Murakami, H. Reevaluation of the D-amino acid
826 compatibility with the elongation event in translation. *J. Am. Chem. Soc.* **135**,
827 1830-1837, (2013).

828 47 Zhu, L. *et al.* Peptide aldehyde inhibitors challenge the substrate specificity of
829 the SARS-coronavirus main protease. *Antiviral Res.* **92**, 204-212, (2011).

830 48 Shi, J. & Song, J. The catalysis of the SARS 3C-like protease is under extensive
831 regulation by its extra domain. *FEBS J.* **273**, 1035-1045, (2006).

832 49 Hendrickson, W. A. Synchrotron crystallography. *Trends Biochem. Sci.* **25**,
833 637-643, (2000).

834 50 Lee, J. *et al.* Crystallographic structure of wild-type SARS-CoV-2 main protease
835 acyl-enzyme intermediate with physiological C-terminal autoprocessing site.
836 *Nat. Commun.* **11**, 5877, (2020).

837 51 Schrödinger Release 2019-1. Protein Preparation Wizard, (Schrödinger, LLC,
838 New York, NY, 2019.).

839 52 Derossi, D., Joliot, A. H., Chassaing, G. & Prochiantz, A. The third helix of the
840 antennapedia homeodomain translocates through biological membranes. *J.*
841 *Biol. Chem.* **269**, 10444-10450, (1994).

842 53 Rogers, J. M., Passioura, T. & Suga, H. Nonproteinogenic deep mutational
843 scanning of linear and cyclic peptides. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 10959-
844 10964, (2018).

845 54 Bashiruddin, N. K. *et al.* Development of cyclic peptides with potent in vivo
846 osteogenic activity through RaPID-based affinity maturation. *Proc. Natl. Acad.*
847 *Sci. U.S.A.* **117**, 31070-31077, (2020).

- 55 Wang, S. *et al.* Glycosylation regulates N-terminal proteolysis and activity of the chemokine CCL14. *ACS Chem. Biol.* **16**, 973-981, (2021).
- 56 Hayashi, Y., Morimoto, J. & Suga, H. In vitro selection of anti-Akt2 thioether-macrocyclic peptides leading to isoform-selective inhibitors. *ACS Chem. Biol.* **7**, 607-613, (2012).
- 57 Norman, A. *et al.* Discovery of cyclic peptide ligands to the SARS-CoV-2 spike protein using mRNA Display. *ACS Cent. Sci.* **7**, 1001-1008, (2021).
- 58 Yamagishi, Y. *et al.* Natural product-like macrocyclic N-methyl-peptide inhibitors against a ubiquitin ligase uncovered from a ribosome-expressed de novo library. *Chem. Biol.* **18**, 1562-1570, (2011).
- 59 Aragão, D. *et al.* MX2: a high-flux undulator microfocus beamline serving both the chemical and macromolecular crystallography communities at the Australian Synchrotron. *J. Synchrotron Radiat.* **25**, 885-891, (2018).
- 60 Kabsch, W. XDS. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **66**, 125-132, (2010).
- 61 Evans, P. R. & Murshudov, G. N. How good are my data and what is the resolution? *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **69**, 1204-1214, (2013).
- 62 McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658-674, (2007).
- 63 Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **66**, 486-501, (2010).
- 64 Murshudov, G. N. *et al.* REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **67**, 355-367, (2011).
- 65 Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **53**, 240-255, (1997).
- 66 Tea, F. *et al.* SARS-CoV-2 neutralizing antibodies: longevity, breadth, and evasion by emerging viral variants. *PLoS Med.* **18**, e1003656, (2021).
- 67 Walker, G. J. *et al.* SARS-CoV-2 in human milk is inactivated by Holder pasteurisation but not cold storage. *J. Paediatr. Child Health* **56**, 1872-1874, (2020).
- 68 Harney, D. J. *et al.* Proteomics analysis of adipose depots after intermittent fasting reveals visceral fat preservation mechanisms. *Cell Rep* **34**, 108804, (2021).
- 69 Martinez-Huenchullan, S. F. *et al.* Blockade of high-fat diet proteomic phenotypes using exercise as prevention or treatment. *Mol. Cell. Proteomics* **20**, 100027, (2020).

Acknowledgements

We acknowledge the ARC Centre of Excellence for Innovations in Peptide and Protein Science (CE200100012 to G.O., C.J.J. and R.J.P.) for funding and the John A. Lamberton Research Scholarship (to J.J-L) and Research Training Program from the Australian government (to J.J-L., S.E.F. and M.J.B) for PhD scholarships. C.N.

acknowledges ARC funding (DE190100015, DP200100348). This research was undertaken in part using cyclic peptide display screening facilities in Sydney Analytical, and the MX2 beamline at the Australian Synchrotron, part of ANSTO, and made use of the Australian Cancer Research Foundation (ACRF) detector. Next generation sequencing was conducted through the Ramaciotti Centre for Genomics. M.L. is a Cancer Institute New South Wales Future Research Leader Fellow. We thank SydneyMS for providing the instrumentation used in this study. We also acknowledge Karishma Patel for her assistance in the use of SEC-MALLS.

Author contributions

J.J-L., G.O., S.T., C.J.J. C.N. and R.J.P conceived the project. J.J-L. performed protein cross-linking and SEC-MALLS. J.J-L. and T.P. performed RaPID selection experiments. J.J-L. S.E.F. and M.J.B performed peptide synthesis. S.U. performed all M^{pro} inhibition assays. K.B.E and M.C.E. produced wild-type and mutant variants of SARS-CoV-2 M^{pro} and performed NMR assignments. R.F. performed M^{pro} crystallographic studies. V.M.S. performed molecular dynamics simulations. A.A. performed all SARS-CoV-2 inhibition assays. A.S.A assessed compounds for cytotoxicity and conducted cellular uptake experiments. M.L. performed proteomics experiments. J.J-L., G.O., S.T., C.J.J. C.N. and R.J.P wrote the paper with contributions from all authors.

Competing interests

The authors declare no competing interests.