



# Mass spectrometric assays monitoring the deubiquitinase activity of the SARS-CoV-2 papain-like protease inform on the basis of substrate selectivity and have utility for substrate identification

Lennart Brewitz<sup>a,b,\*</sup>, H.T. Henry Chan<sup>a</sup>, Petra Lukacik<sup>c,d</sup>, Claire Strain-Damerell<sup>c,d</sup>, Martin A. Walsh<sup>c,d</sup>, Fernanda Duarte<sup>a</sup>, Christopher J. Schofield<sup>a,b,\*</sup>

<sup>a</sup> Chemistry Research Laboratory, Department of Chemistry, University of Oxford, 12 Mansfield Road, OX1 3TA Oxford, United Kingdom

<sup>b</sup> The Ineos Oxford Institute for Antimicrobial Research, Department of Chemistry, University of Oxford, 12 Mansfield Road, OX1 3TA Oxford, United Kingdom

<sup>c</sup> Diamond Light Source Ltd., Harwell Science and Innovation Campus, OX11 0DE Didcot, United Kingdom

<sup>d</sup> Research Complex at Harwell, Harwell Science and Innovation Campus, OX11 0FA Didcot, United Kingdom

## ARTICLE INFO

### Keywords:

SARS-CoV-2 papain-like protease / PL<sup>pro</sup>  
Nucleophilic cysteine protease  
Deubiquitinase / DUB  
Ubiquitin  
Interferon stimulated gene 15 / ISG15  
Mass spectrometry  
Ubiquitin-like modifier  
Virus-host interactions  
Interferon regulatory factor 3 / IRF3  
eIF4E-homologous protein / 4EHP

## ABSTRACT

The SARS-CoV-2 papain-like protease (PL<sup>pro</sup>) and main protease (M<sup>pro</sup>) are nucleophilic cysteine enzymes that catalyze hydrolysis of the viral polyproteins pp1a/1ab. By contrast with M<sup>pro</sup>, PL<sup>pro</sup> is also a deubiquitinase (DUB) that accepts post-translationally modified human proteins as substrates. Here we report studies on the DUB activity of PL<sup>pro</sup> using synthetic N<sup>ε</sup>-lysine-branched oligopeptides as substrates that mimic post-translational protein modifications by ubiquitin (Ub) or Ub-like modifiers (UBLs), such as interferon stimulated gene 15 (ISG15). Mass spectrometry (MS)-based assays confirm the DUB activity of isolated recombinant PL<sup>pro</sup>. They reveal that the sequence of both the peptide fragment derived from the post-translationally modified protein and that derived from the UBL affects PL<sup>pro</sup> catalysis; the nature of substrate binding in the S sites appears to be more important for catalytic efficiency than binding in the S' sites. Importantly, the results reflect the reported cellular substrate selectivity of PL<sup>pro</sup>, i.e. human proteins conjugated to ISG15 are better substrates than those conjugated to Ub or other UBLs. The combined experimental and modelling results imply that PL<sup>pro</sup> catalysis is affected not only by the identity of the substrate residues binding in the S and S' sites, but also by the substrate fold and the conformational dynamics of the blocking loop 2 of the PL<sup>pro</sup>:substrate complex. N<sup>ε</sup>-Lysine-branched oligopeptides thus have potential to help the identification of PL<sup>pro</sup> substrates. More generally, the results imply that MS-based assays with N<sup>ε</sup>-lysine-branched oligopeptides have potential to monitor catalysis by human DUBs and hence to inform on their substrate preferences.

## 1. Introduction

The SARS-CoV-2 genome encodes two nucleophilic cysteine proteases that catalyze hydrolysis of the viral polyproteins pp1a/1ab into functional non-structural proteins (nsps), i.e. the papain-like protease (PL<sup>pro</sup>, a domain of nsp3) and the main protease (M<sup>pro</sup>, nsp5).<sup>1,2</sup> PL<sup>pro</sup> catalyzes the hydrolysis of peptide bonds in pp1a/1ab C-terminal to three LXGG motifs (X represents a non-conserved residue) to release nsps1-3. M<sup>pro</sup> catalyzes the hydrolysis of peptide bonds in pp1a/1ab C-terminal to eleven glutamine residues in (L/F/V)Q(S/A/N) motifs to release nsps 4-16.<sup>1</sup> Catalysis by the two SARS-CoV-2 proteases is essential for viral replication and their inhibition is of therapeutic

relevance,<sup>3-8</sup> i.e. the small-molecules nirmatrelvir and ensitrelvir, which are clinically used to treat COVID-19, inhibit M<sup>pro</sup>.<sup>9,10</sup>

The proteolytic functions of both SARS-CoV-2 M<sup>pro</sup> and, in particular, PL<sup>pro</sup> extend beyond catalyzing the hydrolysis of pp1a/1ab to catalyzing the hydrolysis of (iso)peptide amide bonds in human proteins, in accord with the reported substrate promiscuity of other coronavirus proteases.<sup>11-15</sup> Interestingly, only a few of the human proteins that contain (L/F/V)Q motifs in their coding sequence are validated substrates of isolated recombinant SARS-CoV-2 M<sup>pro</sup>, e.g. human nuclear factor (NF)-κB essential modulator (NEMO), the M<sup>pro</sup>-catalyzed hydrolysis of which has been reported to induce the death of brain endothelial cells.<sup>16,17</sup>

\* Corresponding authors.

E-mail addresses: [lennart.brewitz@chem.ox.ac.uk](mailto:lennart.brewitz@chem.ox.ac.uk) (L. Brewitz), [christopher.schofield@chem.ox.ac.uk](mailto:christopher.schofield@chem.ox.ac.uk) (C.J. Schofield).

<https://doi.org/10.1016/j.bmc.2023.117498>

Received 21 August 2023; Received in revised form 7 October 2023; Accepted 10 October 2023

Available online 12 October 2023

0968-0896/© 2023 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

PL<sup>pro</sup> is proposed to catalyze the hydrolysis of peptide bonds C-terminal to LXGG motifs in the coding sequence of several human proteins, e.g. the serine/threonine-protein kinase ULK1 and protein S, on the basis of *in vitro* and/or cellular studies.<sup>18–20</sup> Although proteomic studies with samples from COVID-19 patients indicate that PL<sup>pro</sup> may accept human proteins as substrates *in vivo*,<sup>21</sup> both the extent and the clinical significance of the PL<sup>pro</sup>-catalyzed degradation of human host proteins during SARS-CoV-2 infections are incompletely understood. Such knowledge is important with respect to understanding the consequences of therapeutic PL<sup>pro</sup> inhibition.

LXGG motifs are conserved in human proteins which are post-translationally modified with ubiquitin (Ub) or ubiquitin-like modifiers (UBL), e.g. interferon stimulated gene 15 (ISG15),<sup>22–25</sup> neural precursor cell expressed, developmentally down-regulated 8 (NEDD8),<sup>26,27</sup> and ubiquitin-related modifier-1 (URM1).<sup>28,29</sup> Ub and UBLs are linked to one or multiple N<sup>ε</sup>-amino groups of a protein lysine residue via an isopeptide amide bond formed with their C-terminal residue, which forms a LXGG motif in the case of Ub, ISG15, NEDD8, and URM1.<sup>30,31</sup> The post-translational (poly)ubiquitylation of human proteins has multiple regulatory functions, including signaling proteins for degradation via the proteasome.<sup>32–34</sup> Protein (poly)ubiquitylation is often reversible; formation of the isopeptide amide bond is catalyzed by Ub/UBL ligases while its hydrolysis is catalyzed by deubiquitinases (DUBs); both Ub/UBL ligases and DUBs may show high levels of specificity, both with respect to the substrate protein and/or the Ub/UBL.<sup>33–38</sup>

SARS-CoV-2 PL<sup>pro</sup> not only catalyzes hydrolysis of the pp1a/1ab polyproteins, but also possesses DUB activity and catalyzes the hydrolysis of isopeptide amide bonds in human proteins which are post-translationally modified with Ub/UBLs,<sup>39,40</sup> similar to the reported DUB activities of PL<sup>pro</sup> from other coronaviruses.<sup>11–13,41</sup> Cellular studies reveal that SARS-CoV-2 PL<sup>pro</sup> preferentially catalyzes the hydrolysis of isopeptide amide bonds C-terminal to ISG15, whereas SARS-CoV PL<sup>pro</sup> prefers diubiquitin (Ub<sub>2</sub>) substrates over ISG15;<sup>39,40</sup> both Ub<sub>2</sub> and ISG15 bind to the active site of PL<sup>pro</sup> and, additionally, to allosteric sites.<sup>42–44</sup> The apparent selectivity of SARS-CoV-2 PL<sup>pro</sup> for catalyzing deISG15ylations is preceded by the substrate selectivity of the human DUB ubiquitin specific peptidase 18 (USP18) which also employs a nucleophilic cysteine to selectively catalyze protein deISG15ylations over protein deubiquitylations.<sup>45</sup> Notably, the DUB activity of SARS-CoV-2 PL<sup>pro</sup> is reported to attenuate the host innate immune response via catalyzing the deISG15ylation of ISG15ylated interferon regulatory factor 3 (IRF3), thus potentially enhancing virulence.<sup>39,46</sup>

Reported assays which monitor the DUB activity of PL<sup>pro</sup> typically

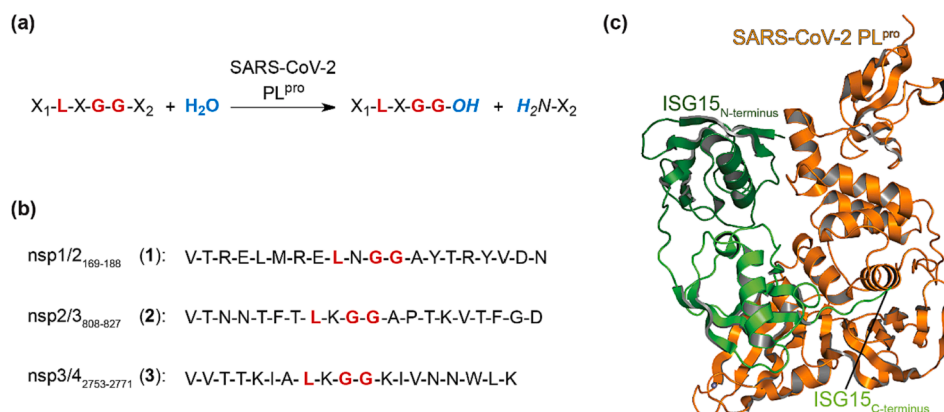
employ Ub/UBL derivatives conjugated to a fluorescent group via a C-terminal amide (e.g. Ub-AMC or ISG15-AMC; AMC: 7-amino-4-methylcoumarin), the hydrolysis of which can be spectroscopically monitored.<sup>39,40,48,49</sup> Matrix-assisted laser desorption/ionization (MALDI) mass spectrometry (MS)-based assays, originally developed to monitor the activity of human DUBs,<sup>50</sup> have also been employed to monitor the DUB activity of PL<sup>pro</sup>; however, the MALDI-MS assay uses Ub<sub>2</sub> as substrate which is apparently not the preferred PL<sup>pro</sup> substrate in cells.<sup>51</sup> The use of substrates which do not bind to the PL<sup>pro</sup> S' sites, the relatively high costs associated with use of the Ub-AMC, ISG15-AMC, and Ub<sub>2</sub> substrates, and/or the use of artificial C-terminal fluorophores limit potential applications of these assays.

We have reported solid-phase extraction coupled to mass spectrometry (SPE-MS)-based assays which directly monitor PL<sup>pro</sup>-catalyzed hydrolysis of oligopeptide fragments derived from the sequence of the Wuhan-Hu-1 strain<sup>52</sup> of SARS-CoV-2 pp1a/1ab (Fig. 1b). The assay results revealed that PL<sup>pro</sup> preferentially catalyzes hydrolysis of an oligopeptide which was based on the LXGG motif separating nsp2 and nsp3, *i.e.* nsp2/3<sub>808–827</sub> (2; Fig. 1).<sup>47</sup> Here we report the results of SPE-MS assays employing N<sup>ε</sup>-lysine branched synthetic oligopeptides as substrates, that bind to both the S and S' sites of PL<sup>pro</sup>, to monitor the DUB activity of isolated recombinant SARS-CoV-2 PL<sup>pro</sup>. The results inform on the substrate requirements for efficient PL<sup>pro</sup> DUB catalysis and reveal that not only the presence of an LXGG motif but also the entire coding sequence of the substrate which binds in proximity of the active site, in particular that of the UBL, determines the catalytic efficiency of PL<sup>pro</sup>.

## 2. Results and discussion

### 2.1. Assaying the DUB activity of SARS-CoV-2 PL<sup>pro</sup> using mass spectrometry

SARS-CoV-2 PL<sup>pro</sup> is reported to attenuate the human innate immune system by catalyzing the deISG15ylation of post-translationally ISG15ylated IRF3;<sup>39</sup> however, to our knowledge, there are no reports on validation of the DUB activity of isolated recombinant PL<sup>pro</sup> with oligopeptide substrates. We thus synthesized SPE-MS-compatible N<sup>ε</sup>-lysine-branched oligopeptides based on the three reported ISG15ylation sites of IRF3, *i.e.* K193, K360, and K366,<sup>53</sup> using solid phase peptide synthesis (SPPS). The IRF3-derived peptide fragment was synthesized on a Rink amide resin from the C- to the N-terminus employing Fmoc strategy, N-Fmoc lysine with a 4-methyltrityl (Mtt) protected N<sup>ε</sup>-amine at the site of the lysine branching, and an N-terminal amino acid with a Boc-protected N<sup>ε</sup>-amine. Branching was introduced via selective



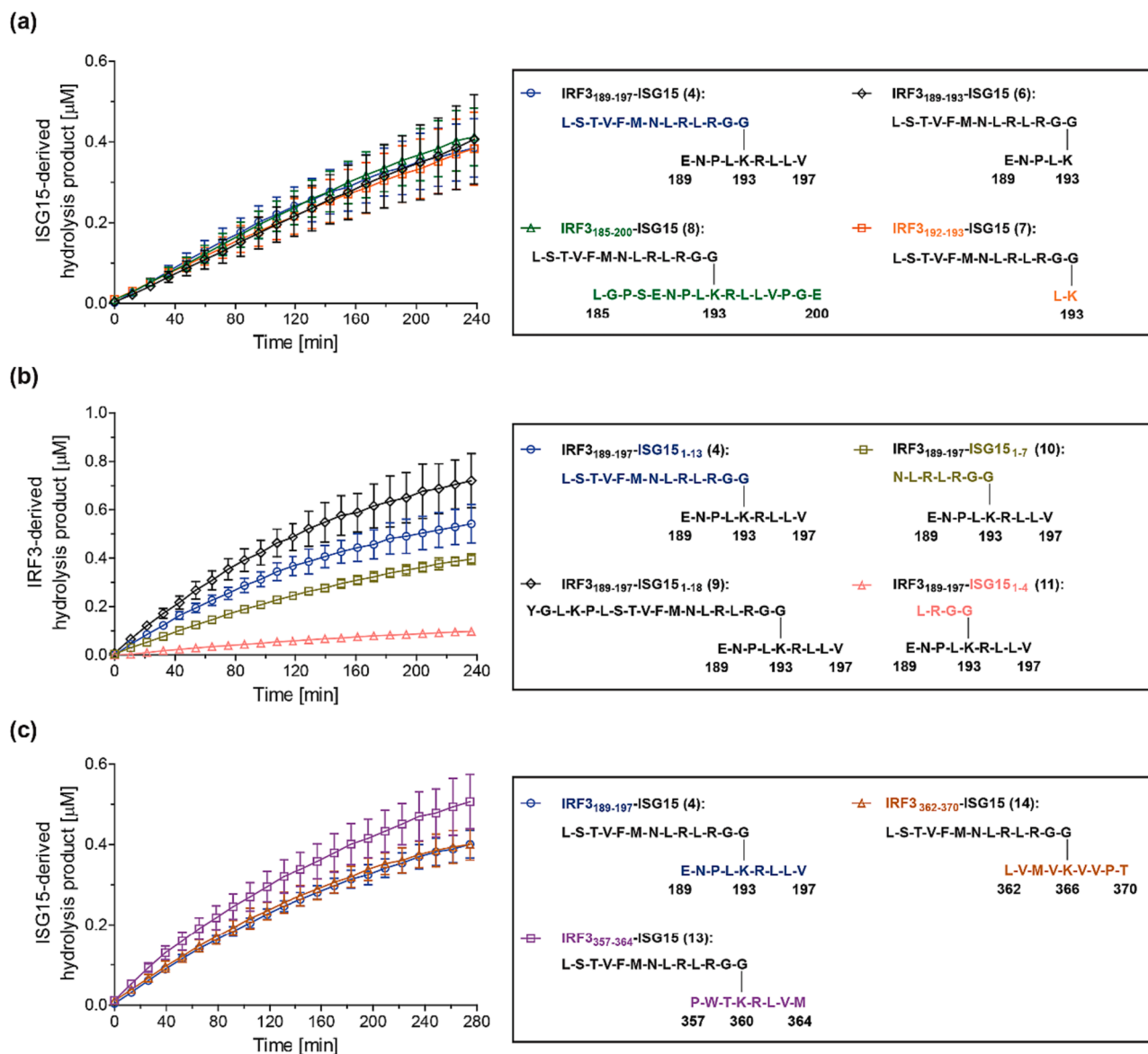
**Fig. 1. PL<sup>pro</sup> catalysis.** (a) PL<sup>pro</sup> catalyzes substrate hydrolysis C-terminal to LXGG motifs; (b) sequences of the reported SARS-CoV-2 pp1a/1ab-derived linear oligopeptides nsp1/2<sub>169–188</sub> (1), nsp2/3<sub>808–827</sub> (2), and nsp3/4<sub>2753–2771</sub> (3), previously used in SPE-MS assays; PL<sup>pro</sup> catalyzed the hydrolysis of 2 substantially more efficiently than that of 1 and 3.<sup>47</sup> Residues of the LXGG motif are in red; (c) SARS-CoV-2 PL<sup>pro</sup> (C111S variant, orange) in complex with ISG15 (PDB ID: 7RBS<sup>44</sup>), which is composed of a C- and N-terminal Ub-like domain (light and dark green, respectively), that indicates binding to both the PL<sup>pro</sup> active site via the C-terminus and the C-terminal Ub-like domain of ISG15 and binding to allosteric sites of PL<sup>pro</sup> via the N-terminal Ub-like domain (dark green).

deprotection of the lysine  $N^{\epsilon}$ -Mtt group using 1% $_{v/v}$  trifluoroacetic acid (TFA) and 2% $_{v/v}$  triisopropylsilane (TIPS) in dichloromethane, followed by SPPS to build the ISG15 fragment from C- to N-termini starting with the  $N^{\epsilon}$ -amino group of the IRF3 lysine at the C-terminus. The purified branched oligopeptides were obtained following global deprotection and cleavage from the resin with TFA and HPLC purification (Supporting Figure S1).

The K<sub>193</sub>-branched IRF3<sub>189-197</sub>-ISG15 oligopeptide **4** was incubated with the isolated recombinant PL<sup>PRO</sup> domain of SARS-CoV-2 nsp3 (0.2  $\mu$ M; enzyme/substrate ratio: 1/10) employing our reported conditions for the PL<sup>PRO</sup>-catalyzed hydrolysis of the linear pp1a/1ab-derived oligopeptide nsp2/3<sub>808-827</sub> (**2**) (50 mM Tris, pH 8.0, 37 °C).<sup>47</sup> Analysis of the reaction mixture after 14 h incubation by SPE-MS indicated apparent

quantitative isopeptide amide bond hydrolysis, as supported by the observed masses of the resultant N- and C-terminal product fragments. Thus, the results confirm the DUB activity of PL<sup>PRO</sup>, which has been assigned on the basis of cellular studies and on studies with ISG15 containing a C-terminal fluorophore,<sup>39,40,48</sup> with an  $N^{\epsilon}$ -lysine-branched oligopeptide substrate.

Time course data were subsequently recorded under the same conditions using an automated SPE-MS setup, however, at ambient temperature rather than at 37 °C. Incomplete levels of PL<sup>PRO</sup>-catalyzed isopeptide amide bond hydrolysis were observed within ~240 min incubations under these conditions (~20 %; Fig. 2a), which may, at least in part, reflect the inability of **4** to efficiently bind to PL<sup>PRO</sup> at allosteric sites as reported for full-length folded (poly)Ub/UBLs.<sup>42-44</sup> The reaction



**Fig. 2.** PL<sup>PRO</sup> catalyzes the hydrolysis of isopeptide amide bonds in  $N^{\epsilon}$ -lysine-branched oligopeptides which mimic ISG15ylated IRF3. (a) Levels of PL<sup>PRO</sup>-catalyzed hydrolysis of the K<sub>193</sub>-branched oligopeptides IRF3<sub>189-197</sub>-ISG15 (**4**; blue circles), IRF3<sub>189-193</sub>-ISG15 (**6**; black diamonds), IRF3<sub>192-193</sub>-ISG15 (**7**; orange boxes), and IRF3<sub>185-200</sub>-ISG15 (**8**; green triangles); (b) levels of PL<sup>PRO</sup>-catalyzed hydrolysis of the K<sub>193</sub>-branched oligopeptides bearing ISG15<sub>1-18</sub> (**9**; black diamonds), ISG15<sub>1-13</sub> (**4**; blue circles), ISG15<sub>1-7</sub> (**10**; olive boxes), and ISG15<sub>1-4</sub> (**11**; salmon triangles); (c) levels of PL<sup>PRO</sup>-catalyzed hydrolysis of the K<sub>360</sub>-branched oligopeptide IRF3<sub>357-364</sub>-ISG15 (**13**; lavender boxes) and the K<sub>366</sub>-branched oligopeptide IRF3<sub>362-370</sub>-ISG15 (**14**; brown triangles) compared to that of the K<sub>193</sub>-branched oligopeptide IRF3<sub>189-197</sub>-ISG15 (**4**; blue circles). Conditions: SARS-CoV-2 PL<sup>PRO</sup> (0.2  $\mu$ M), substrate peptide (2.0  $\mu$ M), inert  $N$ -acetylated inert standard peptide (0.2  $\mu$ M; Ac-LSTVFMNRLRLRG-NH<sub>2</sub> (**5**) in (a) and (c), Ac-ENPLKRLV-NH<sub>2</sub> (**12**) in (b)) in buffer (50 mM Tris, pH 8.0, ambient temperature). Measurement times were normalized to the first sample injection analyzed after the addition of PL<sup>PRO</sup> to the Substrate Mixture ( $t = 0$  s), by which time low levels of substrate hydrolysis were manifest. The inert  $N$ -acetylated hydrolysis products **5** or **12** were used as an internal standard to quantify hydrolysis (Supporting Figures S2 and S3); SPE-MS assay results are a mean of independent triplicates ( $n = 3$ ; mean  $\pm$  standard deviation, SD).

conditions were subsequently varied, with respect to, e.g., buffer composition, pH, and salt additives; the tested variations did not, however, increase the rates of isopeptide amide bond hydrolysis substantially. Note that the corresponding N-terminally *N*-acetylated ISG15-derived hydrolysis product peptide, i.e. Ac-LSTVFMNLRGG-NH<sub>2</sub> (**5**), was added as an inert internal standard to the reaction mixture to enable quantification of PL<sup>pro</sup>-catalyzed product formation; the presence of **5** in the reaction mixture did not affect PL<sup>pro</sup> catalysis, at least substantially (Supporting Figure S2).

Derivatives of the K<sub>193</sub>-branched IRF3<sub>189-197</sub>-ISG15 oligopeptide **4** were synthesized to investigate whether varying the length of the IRF3- or the ISG15-derived fragments affects PL<sup>pro</sup> catalysis (Fig. 2a and 2b). Interestingly, SPE-MS analysis of the PL<sup>pro</sup>-catalyzed hydrolysis of the oligopeptides **4**, **6**, **7**, and **8** revealed that varying the length of the IRF3-derived fragment did not affect catalysis substantially, though the hydrolysis of the K<sub>193</sub>-branched IRF3<sub>185-200</sub>-ISG15 oligopeptide **8**, which has the relatively largest IRF3-derived fragment of the tested oligopeptides, appeared to be slightly favored (Fig. 2a). By contrast, varying the length of the ISG15-derived fragment had a substantial effect on PL<sup>pro</sup> catalysis; the oligopeptide **9** with the relatively largest ISG15-derived fragment bearing 18 C-terminal amino acids was hydrolyzed substantially more efficiently than oligopeptide **11** with the relatively smallest ISG15-derived fragment bearing only the essential C-terminal LRGG tetrad (Fig. 2b). Levels of PL<sup>pro</sup>-catalyzed substrate hydrolysis appear to increase with increasing length of the ISG15-derived fragment peptide; the rank order of PL<sup>pro</sup> substrate preference was determined by quantifying product formation relative to Ac-ENPLKRLLV-NH<sub>2</sub> (**12**), which was used as an internal standard in the reactions. The combined oligopeptide results indicate that the binding of the N<sup>ε</sup>-lysine-branched oligopeptides to the S' sites is less important for catalysis than binding to the S sites. The situation may, however, be different with full-length folded protein substrates. Note that either the corresponding N-terminally *N*-acetylated ISG15-derived hydrolysis product peptide, i.e. Ac-LSTVFMNLRGG-NH<sub>2</sub> (**5**), or the corresponding N-terminally *N*-acetylated IRF3<sub>189-197</sub>-derived hydrolysis product peptide, i.e. Ac-ENPLKRLLV-NH<sub>2</sub> (**12**), was added as an internal standard to the reaction mixtures to enable quantification of PL<sup>pro</sup>-catalyzed product formation; the presence of **5** or **12** in the reaction mixture did not affect PL<sup>pro</sup> catalysis, at least substantially (Supporting Figures S2 and S3).

The K<sub>360</sub>-branched IRF3<sub>357-364</sub>-ISG15 peptide **13** and the K<sub>366</sub>-branched IRF3<sub>362-370</sub>-ISG15 peptide **14** were synthesized and tested to investigate the potential of isolated recombinant PL<sup>pro</sup> to catalyze regioselective deISG15ylations of post-translationally modified IRF3. Oligopeptides **4**, **13**, and **14** were incubated with PL<sup>pro</sup> in the presence of the internal standard **5**, and isopeptide amide bond hydrolysis was monitored using SPE-MS (Fig. 2c). The results reveal that the sequences of the three tested IRF3-derived peptide fragments, which likely bind to the S' sites of PL<sup>pro</sup>, do not affect the rate of peptide hydrolysis substantially, in accord with prior results which showed that the length of the IRF3-derived peptide fragment affects PL<sup>pro</sup> catalysis less substantially than the length of the ISG15-derived peptide fragment (Fig. 2a and 2b). The observation that PL<sup>pro</sup> may catalyze the deISG15ylation of oligopeptide **13** slightly more efficiently than that of **4** and **14** might be a result of the comparatively lower purity of **13** compared to that of all other tested peptides (Supporting Figure S1). Note that PL<sup>pro</sup> may catalyze the regioselective deISG15ylation of full-length folded IRF3 in a cellular context.

## 2.2. The primary sequence of the PL<sup>pro</sup> substrates affects catalysis

To investigate whether the substrate preference of SARS-CoV-2 PL<sup>pro</sup> depends on the primary sequence of the UBLs, we synthesized the oligopeptide derivatives of K<sub>193</sub>-branched IRF3<sub>189-197</sub>-ISG15 **4**, in which the ISG15-fragment was substituted for the corresponding first 13 C-terminal amino acids of Ub (IRF3<sub>189-197</sub>-Ub, **15**) or the UBLs NEDD8 (IRF3<sub>189-197</sub>-NEDD8, **16**), URM1 (IRF3<sub>189-197</sub>-URM1, **17**), and small

ubiquitin-related modifier 1 (SUMO1; IRF3<sub>189-197</sub>-SUMO1, **18**), the latter of which does not contain a C-terminal LXGG motif and should hence not be a substrate of PL<sup>pro</sup> (Supporting Figure S1). K<sub>193</sub>-branched IRF3<sub>189-197</sub> derivatives were attractive synthesis targets because the ubiquitylation at K193 had been reported<sup>54</sup>; the corresponding NEDD8, URM1, and SUMO1 derivatives of **4** were also synthesized despite their unclear biological relevance. The length of the UBL fragment was kept constant to help enable comparison, as PL<sup>pro</sup> catalyzes the hydrolysis of derivatives of **4**, in which the length of the ISG15 fragment was varied, with different efficiencies (Fig. 2b).

The oligopeptides **4**, **15**, **16**, **17**, and **18** were incubated with isolated recombinant SARS-CoV-2 PL<sup>pro</sup>; hydrolysis of their isopeptide amide bond was monitored in the presence of the corresponding inert N-terminally *N*-acetylated IRF3<sub>189-197</sub>-derived hydrolysis product peptide Ac-ENPLKRLLV-NH<sub>2</sub> (**12**) using SPE-MS (Fig. 3a). The results revealed that PL<sup>pro</sup> catalyzes the hydrolysis of the IRF3<sub>189-197</sub>-ISG15 (**4**) isopeptide amide bond more efficiently than those of IRF3<sub>189-197</sub>-Ub (**15**) and, in particular, IRF3<sub>189-197</sub>-NEDD8 (**16**) (Fig. 3a). However, **4**, **15**, and **16** were substantially more efficient PL<sup>pro</sup> substrates than IRF3<sub>189-197</sub>-URM1 (**17**). As anticipated, IRF3<sub>189-197</sub>-SUMO1 (**18**) was not a substrate for isolated recombinant PL<sup>pro</sup> under the tested conditions, an observation which likely reflects the lack of an LXGG motif in **18**, which instead has a QTGG motif that cannot interact efficiently with the S1-S4 sites of PL<sup>pro</sup> (*vide infra*).

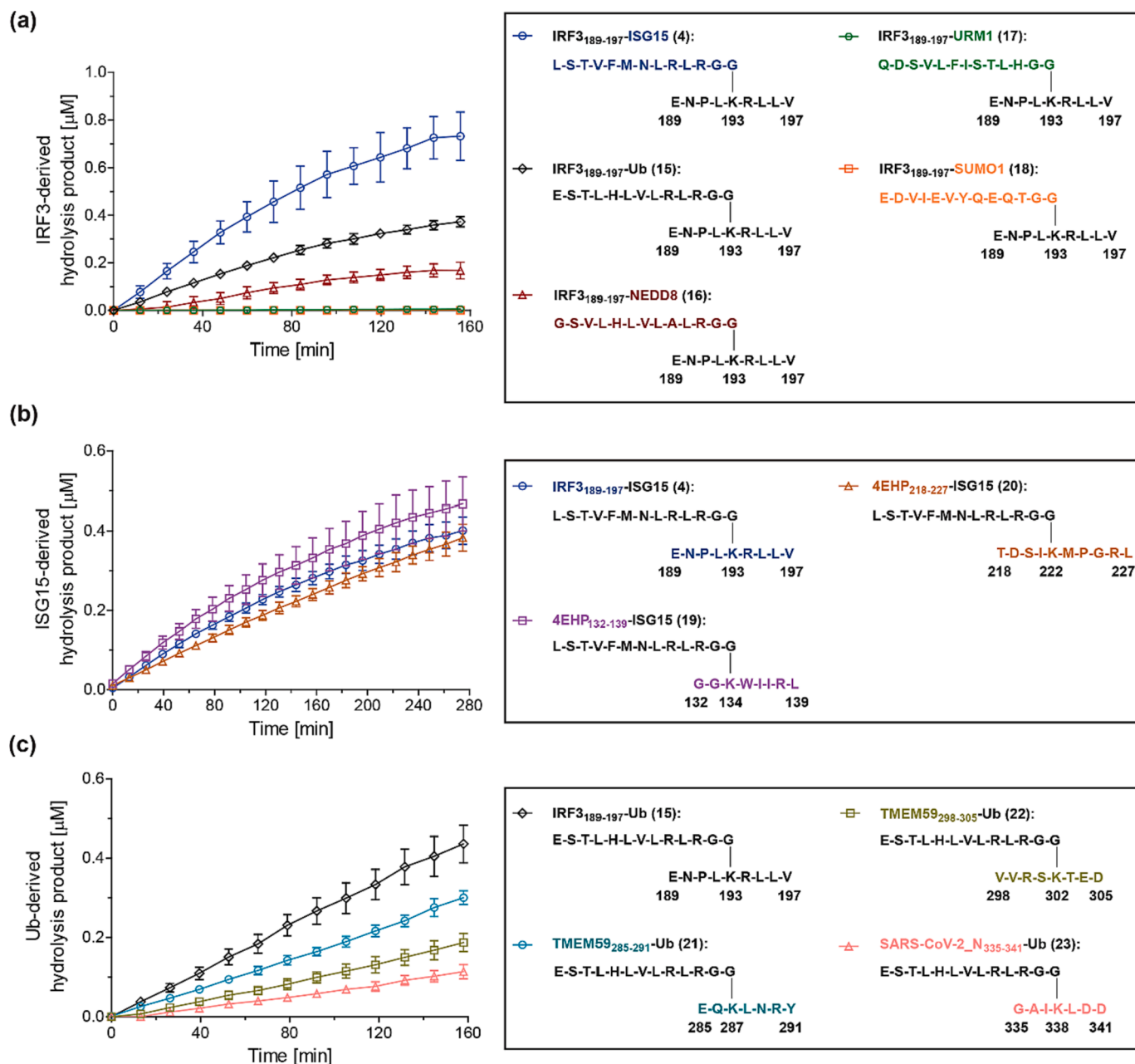
IRF3<sub>189-197</sub>-ISG15 (**4**), IRF3<sub>189-197</sub>-Ub (**15**), and IRF3<sub>189-197</sub>-NEDD8 (**16**) all have an LRGG motif, whereas IRF3<sub>189-197</sub>-URM1 (**17**) has an LHGG motif which may rationalize why it is a substantially poorer substrate than **4**, **15**, and **16** (Fig. 3a). Nonetheless, the assay results clearly show that the different hydrolysis rates of **4**, **15**, and **16** are not dependent on the presence of a LRGG motif or an IRF3<sub>189-197</sub> fragment, suggesting that the primary/coding sequence of the UBL N-terminal to the LXGG motif is important in PL<sup>pro</sup> catalysis.

Interestingly, the results with **4**, **15**, and **16** reflect reported SARS-CoV-2 PL<sup>pro</sup> substrate preferences observed in cellular studies which also showed that PL<sup>pro</sup>-catalyzed protein deISG15ylations are more efficient than the corresponding deubiquitylations, which are both more efficient than deNEDD8ylations.<sup>39,40</sup> Thus, the combined evidence suggests that not only the substrate fold affects PL<sup>pro</sup> catalysis, *inter alia* via binding to allosteric sites, but also, in particular, the sequence identity of the amino acid fragment N-terminal to the LXGG motif of the substrate.

In general, the ISG15ylation sites of human proteins are poorly characterized, in part, likely because the standard workflow to localize protein ubiquitylation sites, i.e. protein denaturation followed by tryptic digestion and analysis of the resultant peptide fragments by MS/MS, affords the same lysine-branched peptide fragments for both ubiquitylated and ISG15ylated proteins. Thus, a different workflow is required, e.g. involving digestion using non-standard proteases with different substrate selectivities than trypsin.<sup>35,55–58</sup> Apart from ISG15ylation sites of IRF3, the ISG15ylation sites of the eIF4E-homologous protein (4EHP) have been characterized, i.e. at K134 and K222.<sup>59</sup> 4EHP binds to the mRNA 5'-cap structure and suppresses translation by competing with the eukaryotic initiation factor 4E (eIF4E) for binding to the cap structure. The ISG15ylation of 4EHP is proposed to control translation during immune responses;<sup>59</sup> note that the translation initiation factor eIF4G is a reported substrate of rhinovirus 2A<sup>pro</sup> and of the foot-and-mouth-disease virus (FMDV) leader protease (L<sup>pro</sup>), suggesting that modulation of protein translation may be a general strategy of *Riboviria* to evade the host immune system.<sup>60</sup>

The K<sub>134</sub>-branched 4EHP<sub>132-139</sub>-ISG15 peptide **19** and the K<sub>222</sub>-branched 4EHP<sub>218-227</sub>-ISG15 peptide **20** were synthesized to investigate the potential of isolated recombinant SARS-CoV-2 PL<sup>pro</sup> to catalyze selective deISG15ylations of post-translationally modified 4EHP. Oligopeptides **4**, **19**, and **20** were incubated with PL<sup>pro</sup> in the presence of the inert internal standard **5**, and isopeptide amide bond hydrolysis was monitored using SPE-MS (Fig. 3b). The results reveal that the rates of





**Fig. 3.** PL<sup>pro</sup> catalyzes the hydrolysis of isopeptide amide bonds in N<sup>ε</sup>-lysine-branched oligopeptides that mimic post-translational protein modification by Ub/UBLs. (a) Levels of PL<sup>pro</sup>-catalyzed hydrolysis of oligopeptides derived from K<sub>193</sub>-branched IRF3<sub>189-197</sub> with ISG15 (4; blue circles), Ub (15; black diamonds), NEDD8 (16; red triangles), URM1 (17; green hexagons), and SUMO1 (18; orange boxes). Note that SUMO1 does not have an LXGG motif at its C-terminus; (b) levels of PL<sup>pro</sup>-catalyzed hydrolysis of the K<sub>134</sub>-branched oligopeptide 4EHP<sub>132-139</sub>-ISG15 (19; lavender boxes) and the K<sub>222</sub>-branched oligopeptide 4EHP<sub>218-227</sub>-ISG15 (20; brown triangles) compared to that of the K<sub>193</sub>-branched oligopeptide IRF3<sub>189-197</sub>-ISG15 (4; blue circles); (c) levels of PL<sup>pro</sup>-catalyzed hydrolysis of oligopeptides derived from the N<sup>ε</sup>-lysine-branched IRF3<sub>189-197</sub>-Ub (15; black diamonds), TMEM59<sub>285-291</sub>-Ub (21; teal circles), TMEM59<sub>298-305</sub>-Ub (22; olive boxes), and SARS-CoV-2<sub>N335-341</sub>-Ub (23; salmon triangles) peptides. Conditions: SARS-CoV-2 PL<sup>pro</sup> (0.2 μM), substrate peptide (2.0 μM), inert N-acetylated standard peptide (0.2 μM; Ac-ENPLKRLLV-NH<sub>2</sub> (12) in (a), Ac-LSTVFMNLRRLRGG-NH<sub>2</sub> (5) in (b), and Ac-ESTLHLVRLRLRGG-NH<sub>2</sub> (24) in (c)) in buffer (50 mM Tris, pH 8.0, ambient temperature). Measurement times were normalized to the first sample injection analyzed after the addition of PL<sup>pro</sup> to the Substrate Mixture (t = 0 s), by which time low levels of substrate hydrolysis were manifest. The inert N-acetylated hydrolysis product 5, 12 or 24 was used as an internal standard to quantify hydrolysis (Supporting Figures S2-S4); SPE-MS assay results are a mean of independent triplicates (n = 3; mean ± SD).

PL<sup>pro</sup>-catalyzed peptide hydrolyses for 4, 19, and 20 are similar, within experimental error (Fig. 3b). This observation is in accord with the previous results showing that PL<sup>pro</sup> catalyzes the deISG15ylation of IRF3-derived fragment peptides regardless of the site of ISG15ylation (Fig. 2c) and thus supports the proposal that the substrate binding in the S sites of PL<sup>pro</sup> is more important for efficient catalysis than substrate binding in the S' sites of PL<sup>pro</sup>, in accord with reported studies on the substrate efficiency of oligopeptides based on SARS-CoV-2 nsps.<sup>61</sup> Note

that the observed variation in the absolute efficiency of PL<sup>pro</sup> catalysis appeared to depend on the batch of PL<sup>pro</sup>; hence, we determined the rank order of substrate preference by quantifying product formation relative to internal standards, the rank order did not depend on the batch of PL<sup>pro</sup>. The results suggest that ISG15ylated 4EHP may be a substrate of PL<sup>pro</sup> in cells.

### 2.3. *N*<sup>ε</sup>-Lysine-branched oligopeptides may help enable the identification of PL<sup>pro</sup> substrates

To further dissect the effect of the UBL fragment of the tested *N*<sup>ε</sup>-lysine-branched oligopeptides, which binds to the PL<sup>pro</sup> S sites, on the rate of isopeptide amide bond hydrolysis from that of the fragment which binds to the PL<sup>pro</sup> S' sites, and to investigate the potential of synthetic *N*<sup>ε</sup>-lysine-branched oligopeptides for the identification of SARS-CoV-2 PL<sup>pro</sup> substrates, we synthesized a set of *N*<sup>ε</sup>-lysine-branched oligopeptides which mimic protein ubiquitinylation, *i.e.* via employing the 13 C-terminal residues of Ub. In general, the sites of protein (poly) ubiquitinylations have been characterized in greater detail than the sites of protein ISG15ylations. Potential SARS-CoV-2 PL<sup>pro</sup> substrates were chosen on the basis of reported proteomic studies which showed, *e.g.*, that K338 of the SARS-CoV-2 nucleocapsid (N) protein can be ubiquitinylated.<sup>62</sup> Reduced levels of ubiquitinylation of the autophagy-related human transmembrane protein TMEM59 (at *e.g.* K287 and K302) have been associated with SARS-CoV-2 infections, which may suggest that TMEM59 could be a substrate of SARS-CoV-2 PL<sup>pro</sup>.<sup>62</sup> We thus synthesized the corresponding K<sub>287</sub>-branched TMEM59<sub>285-291</sub>-Ub peptide **21**, the K<sub>302</sub>-branched TMEM59<sub>298-305</sub>-Ub peptide **22**, and the K<sub>338</sub>-branched SARS-CoV-2\_N<sub>335-341</sub>-Ub peptide **23** (Supporting Figure S1).

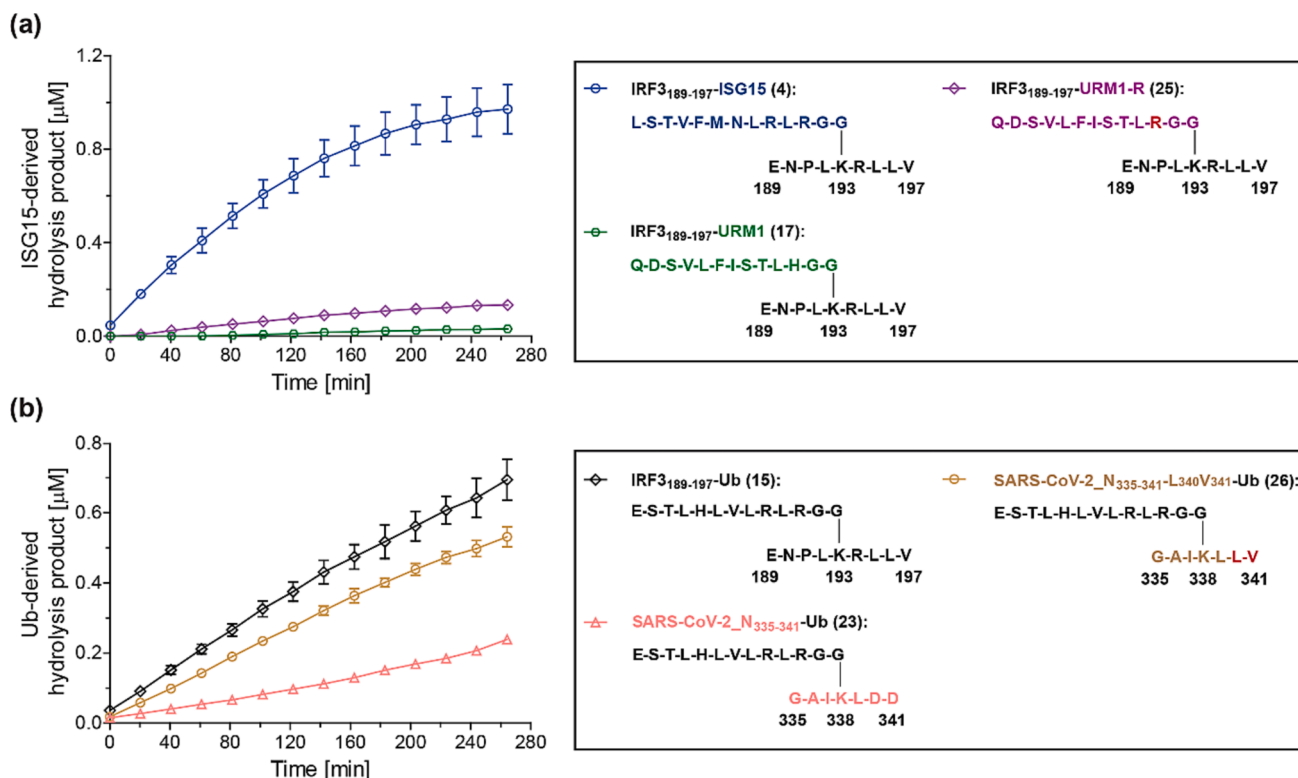
The oligopeptides **15**, **21**, **22**, and **23** were incubated with isolated recombinant SARS-CoV-2 PL<sup>pro</sup> and hydrolysis of their isopeptide amide bond was monitored using SPE-MS (Fig. 3c). Note that the corresponding N-terminally *N*-acetylated Ub-derived hydrolysis product peptide, *i.e.* Ac-ESTLHLVRLRG-NH<sub>2</sub> (**24**; Supporting Figure S1), was added as an inert internal standard to the reaction mixtures to enable

quantification of PL<sup>pro</sup>-catalyzed product formation; the presence of **24** in the reaction mixture did not affect PL<sup>pro</sup> catalysis substantially (Supporting Figure S4). The results reveal that SARS-CoV-2 PL<sup>pro</sup> catalyzes the hydrolysis of the IRF3<sub>189-197</sub>-Ub (**15**) isopeptide amide bond more efficiently than those of **21**, **22**, and **23** (Fig. 3c). The PL<sup>pro</sup>-catalyzed isopeptide hydrolysis in IRF3<sub>189-197</sub>-Ub (**15**) was ~2-fold more efficient as in SARS-CoV-2\_N<sub>335-341</sub>-Ub (**23**), which was the least efficient substrate of the four substrates tested. This observation indicates that PL<sup>pro</sup> catalysis can be affected by substrate binding to the S' sites and, by implication, also by the substrate fold in proximity of the S' sites, by contrast to previous results with peptide fragments mimicking the ISG15ylation of IRF3 and 4EHP (Fig. 2).

The combined results suggest that PL<sup>pro</sup> may be promiscuous (including compared to M<sup>pro</sup>) with respect to its *in vivo* protein substrates, and that it likely has the capability to catalyze a broad range of deISG15ylation and deubiquitinylation reactions of both host and viral proteins. The results also highlight the potential of *N*<sup>ε</sup>-lysine-branched oligopeptides to help enable identification of PL<sup>pro</sup> substrates, including post-translationally modified viral proteins.

### 2.4. Single amino acid substitutions in PL<sup>pro</sup> substrates can affect catalysis

The observation that IRF3<sub>189-197</sub>-ISG15 (**4**), IRF3<sub>189-197</sub>-Ub (**15**), and IRF3<sub>189-197</sub>-NEDD8 (**16**) all have an LRGG motif, whereas IRF3<sub>189-197</sub>-URM1 (**17**) has an LHGG motif and, relative to **4**, **15**, and **16**, is a substantially poorer PL<sup>pro</sup> substrate raises the question as to whether PL<sup>pro</sup> catalyzes the hydrolysis of the corresponding IRF3<sub>189-197</sub>-URM1-R (**25**) variant of **17**, in which the LHGG histidine residue has been substituted for an arginine residue, more efficiently than that of **17** (Fig. 3a). **25** was



**Fig. 4. The primary sequence of *N*<sup>ε</sup>-lysine-branched oligopeptides affects the efficiency of PL<sup>pro</sup> catalysis.** (a) PL<sup>pro</sup>-catalyzed hydrolysis of the H-to-R variant of the K<sub>193</sub>-branched IRF3<sub>189-197</sub>-URM1 peptide **17** (green hexagons), *i.e.* **25** (lavender diamonds), compared to that of IRF3<sub>189-197</sub>-ISG15 (**4**; blue circles); (b) PL<sup>pro</sup>-catalyzed hydrolysis of SARS-CoV-2\_N<sub>335-341</sub>-Ub\_L<sub>340</sub>V<sub>341</sub> (**26**; brown circles) compared to that of SARS-CoV-2\_N<sub>335-341</sub>-Ub (**23**; salmon triangles) and IRF3<sub>189-197</sub>-Ub (**15**; black diamonds). Substituted residues are in red. Conditions: SARS-CoV-2 PL<sup>pro</sup> (0.2 μM), substrate peptide (2.0 μM), inert *N*-acetylated standard peptide (0.2 μM; Ac-ENPLKRLV-NH<sub>2</sub> (**12**) in (a), Ac-ESTLHLVRLRG-NH<sub>2</sub> (**24**) in (b)) in buffer (50 mM Tris, pH 8.0, ambient temperature). Measurement times were normalized to the first sample injection analyzed after the addition of PL<sup>pro</sup> to the Substrate Mixture (*t* = 0 s), by which time low levels of substrate hydrolysis were manifest. The inert *N*-acetylated hydrolysis products were used as internal standard to quantify hydrolysis (Supporting Figures S3 and S4); SPE-MS assay results are a mean of independent triplicates (*n* = 3; mean ± SD).

thus synthesized (Supporting Figure S1) and incubated with isolated recombinant PL<sup>pro</sup>, and its hydrolysis was monitored using SPE-MS (Fig. 4a).

Although the PL<sup>pro</sup>-catalyzed hydrolysis of **25** was ~4-fold more efficient than that of **17** (after 4 h), it remained ~7-fold less efficient compared to the PL<sup>pro</sup>-catalyzed hydrolysis of IRF3<sub>189-197</sub>-ISG15 (**4**) (after 4 h), indicating that peptides with an LRGG motif are, in principle, better substrates than those with an LHGG motif (Fig. 4a). However, considering that **25** and **4** are based on the identical IRF3<sub>189-197</sub> fragment C-terminal to their LRGG motif which likely binds to the S' sites of PL<sup>pro</sup> in an identical manner, the results support the proposal that the primary sequence of the substrate N-terminal to the LXGG motif can substantially affect catalysis. This observation may reflect the importance of substrate binding to S sites other than S1-S4, including at allosteric positions, and/or the presence of secondary structural elements in the substrate that favor PL<sup>pro</sup> catalysis. Note that substrate binding to allosteric sites by full-length folded substrate proteins has been proposed to be important in PL<sup>pro</sup> catalysis on the basis of crystallographic studies.<sup>42-44</sup>

The PL<sup>pro</sup>-catalyzed hydrolysis of TMEM59<sub>298-305</sub>-Ub (**22**) and SARS-CoV-2<sub>N335-341</sub>-Ub (**23**), which bear an acidic ED or DD sequence, respectively, in the sequence C-terminal to their LXGG motif, is less efficient compared to that of IRF3<sub>189-197</sub>-Ub (**15**) (Fig. 3c), which bears a hydrophobic LV sequence at the corresponding position. This observation raises the possibility that acidic residues proximate to the substrate P1' lysine residue involved in the isopeptide amide bond, are detrimental for PL<sup>pro</sup> catalysis. To test this proposal, we synthesized the corresponding SARS-CoV-2<sub>N335-341</sub>-Ub<sub>L340V341</sub> variant (**26**; Supporting Figure S1), in which the DD sequence was substituted for an LV sequence as present in **15**, and monitored its PL<sup>pro</sup>-catalyzed hydrolysis using SPE-MS (Fig. 4b). The results reveal that PL<sup>pro</sup> catalyzes the hydrolysis of **26** with similar efficiency as that of **15** and ~2.5-fold more efficiently than that of **23** (after 4 h), supporting the proposal that hydrophobic residues may be preferred in proximity of the substrate lysine residue involved in the isopeptide amide bond, potentially due to improved binding to the PL<sup>pro</sup> S' sites.

The combined results clearly indicate that the primary sequence of oligopeptides affects SARS-CoV-2 PL<sup>pro</sup> catalysis and that apparently minor changes in their sequence, which may be distal to the LXGG motif, can have a pronounced effect on catalysis. It thus appears that substrate binding to both the S and S' sites of PL<sup>pro</sup> has potential to modulate catalysis.

## 2.5. Substrate competition studies

To investigate whether N<sup>ε</sup>-lysine-branched oligopeptides or the reported pp1a/1ab-derived linear oligopeptide nsp2/3<sub>808-827</sub> (**2**) are more efficient substrates of SARS-CoV-2 PL<sup>pro</sup>, we attempted to determine  $k_{cat}/K_m$ -values. These efforts were, however, unsuccessful due to technical limitations of the SPE-MS assay, i.e. high peptide concentrations saturated the sensor of the mass spectrometer impeding the quantification of PL<sup>pro</sup> catalysis. As an alternative, direct substrate competition studies were performed using the linear pp1a/1ab-derived oligopeptide nsp2/3<sub>808-827</sub> (**2**) and the two lysine-branched oligopeptide PL<sup>pro</sup> substrates IRF3<sub>189-197</sub>-ISG15 (**4**) or IRF3<sub>189-197</sub>-Ub (**15**). Note that nsp2/3<sub>808-827</sub> was employed as a pp1a/1ab-derived linear substrate because it was a more efficient substrate compared to linear peptides based on the nsp1/2 or nsp3/4 cleavage site,<sup>47</sup> in accord with studies that employed a LCMS assay to investigate the substrate preference of SARS-CoV-2 PL<sup>pro</sup>.<sup>63</sup>

Equimolar amounts of nsp2/3<sub>808-827</sub> (**2**) and either IRF3<sub>189-197</sub>-ISG15 (**4**) or IRF3<sub>189-197</sub>-Ub (**15**) were incubated with isolated recombinant SARS-CoV-2 PL<sup>pro</sup> in the same reaction vessel, together with the corresponding inert N-acetylated hydrolysis product peptides of **2** and **4** or **15** (as internal standards). PL<sup>pro</sup>-catalyzed peptide hydrolysis was monitored using SPE-MS, which was suitable for these experiments provided

that substrates and products have different masses.<sup>64</sup> The combined results reveal that PL<sup>pro</sup>-catalyzed hydrolysis of **2** and **4/15** does not affect each other under the tested conditions, i.e. the PL<sup>pro</sup>-catalyzed hydrolysis of **2** proceeds as efficiently in the presence of **4** or **15** as in its absence, and vice versa (Fig. 5). This observation may indicate that, under the tested conditions, the initial binding of the peptides to PL<sup>pro</sup> is not rate-limiting regardless of whether peptides are branched or linear, i.e. the on/off-rates for peptide binding are faster than the peptide hydrolysis rates, potentially because the peptides lack appropriate folding.

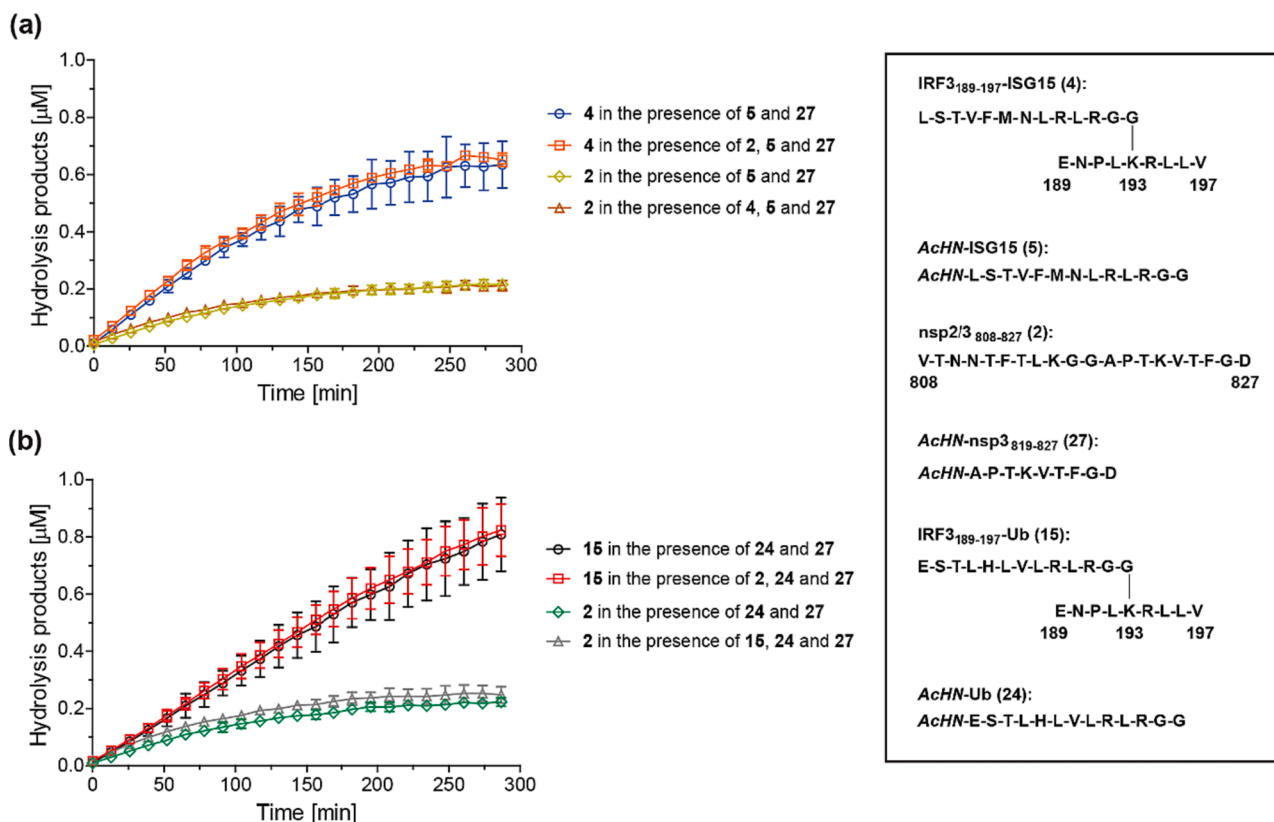
The results also reveal that the PL<sup>pro</sup>-catalyzed hydrolysis of nsp2/3<sub>808-827</sub> (**2**) is less efficient than that of **4** or **15**, i.e. ~10 % of **2** was observed to be hydrolyzed after 5 h whereas ~35–40 % of **4** and **15** were observed to be hydrolyzed after 5 h which may indicate that the  $k_{cat}/K_m$ -values of the N<sup>ε</sup>-lysine-branched oligopeptides are higher than that of **2**. Notably, it appears that the absolute levels of PL<sup>pro</sup> catalysis depended on the batch of PL<sup>pro</sup> and substrate sequence used.

## 2.6. Conformational analysis of substrate binding

To investigate the binding mode of N<sup>ε</sup>-lysine-branched oligopeptides to PL<sup>pro</sup>, we computationally modelled nine non-covalently bound PL<sup>pro</sup>: oligopeptide complexes, including the IRF3<sub>189-197</sub> oligopeptide N<sup>ε</sup>-branched with ISG15 (**4**), Ub (**15**), NEDD8 (**16**), URM1 (**17**) or SUMO1 (**18**), as well as the TMEM59<sub>285-291</sub> (**21**), TMEM59<sub>298-305</sub> (**22**), and SARS-CoV-2<sub>N335-341</sub> (**23**) oligopeptides, all of which are branched with Ub via an N<sup>ε</sup>-lysine. The models were based on a reported PL<sup>pro</sup> structure (PDB ID: 6WX4<sup>65</sup>), using two of the reported AutoDock CrankPep (ADCP)<sup>66</sup>-docked poses of the nsp2/3<sub>808-827</sub> oligopeptide **2**<sup>61</sup> as templates for modelling the N- and C-terminal fragments (Supporting Figures S5-S7). For **17**, we modelled both the neutral (<sup>N</sup>) and positive charged (<sup>+</sup>) states of the P3 histidine imidazole group, as PL<sup>pro</sup> has a reported preference for positively charged residues at this position.<sup>65</sup> The resultant PL<sup>pro</sup>:oligopeptide complexes were subjected to 3 × 200 ns molecular dynamics (MD) simulations, and the stability of the complexes was assessed by calculating the root mean square deviations (RMSDs) and root mean square fluctuations (RMSFs) (Supporting Figures S8-S17).

Analysis of the backbone RMSFs of the PL<sup>pro</sup>-bound N<sup>ε</sup>-lysine-branched oligopeptides reveals that the P6-P1 residues of the UBL-derived peptide fragment are stable, with values <3 Å (Supporting Figure S14); an exception was IRF3<sub>189-197</sub>-URM1 (**17**) with a neutral P3 histidine imidazole, an observation suggesting that the protonation state of this group may affect complex stability. RMSF analysis of the oligopeptide sidechains reveals that the conformation of the P4 residue is rigid relative to its adjacent P5 and P3 residues (Supporting Figure S15), likely reflecting the conserved nature of the LXGG motif leucine residue which binds to the PL<sup>pro</sup> S4 site. Notably, the conformation of the P4 glutamine residue of the IRF3<sub>189-197</sub>-SUMO1 oligopeptide (**18**), which is not a substrate of isolated recombinant PL<sup>pro</sup> and which lacks the LXGG motif (Fig. 3a), is also rigid, suggesting that complex stability itself may not necessarily indicate productive substrate recognition.

Conserved hydrogen bonding interactions appear to stabilize binding of the oligopeptide backbone to the PL<sup>pro</sup> S4-S1 sites (Fig. 6, Supporting Figures S18 and S19), reminiscent of the interactions observed in the modelled PL<sup>pro</sup>:nsp2/3<sub>808-827</sub> (**2**) complex.<sup>61</sup> Hydrogen bonding interactions involving residues of the UBL-derived peptide fragment N-terminal to P4-P1 may also contribute to binding, e.g. with T75 and Q174; however, they appear to be less conserved than those with the P4-P1 LXGG motif (Supporting Table S1, Fig. 6, Supporting Figures S18 and S19), consistent with previous observations on the modelled interactions of pp1a/1ab-derived linear oligopeptides with PL<sup>pro</sup>.<sup>61</sup> Notably, the IRF3<sub>189-197</sub>-SUMO1 oligopeptide (**18**) can apparently engage in similar interactions with PL<sup>pro</sup> as the other modelled N<sup>ε</sup>-lysine-branched oligopeptides; however, it appears that its P4 glutamine sidechain binds to the S4 site less efficiently than a leucine sidechain (Supporting Figures S20-S24).



**Fig. 5.** *N*<sup>ε</sup>-Lysine-branched oligopeptides are more efficient substrates of isolated recombinant PL<sup>pro</sup> than linear oligopeptides. (a) The SARS-CoV-2 PL<sup>pro</sup>-catalyzed hydrolysis of IRF3<sub>189-197</sub>-ISG15 (4) is not affected by the presence of nsp2/3<sub>808-827</sub> (2) and vice versa; 4 appears to be a more efficient PL<sup>pro</sup> substrate than 2; (b) the SARS-CoV-2 PL<sup>pro</sup>-catalyzed hydrolysis of IRF3<sub>189-197</sub>-Ub (15) is not affected by the presence of nsp2/3<sub>808-827</sub> (2) and vice versa; 15 appears to be a more efficient PL<sup>pro</sup> substrate than 2. Conditions: SARS-CoV-2 PL<sup>pro</sup> (0.2  $\mu$ M), substrate peptide(s) (2.0  $\mu$ M), inert *N*-acetylated standard peptides (0.2  $\mu$ M, as indicated) in buffer (50 mM Tris, pH 8.0, ambient temperature). Measurement times were normalized to the first sample injection analyzed after the addition of PL<sup>pro</sup> to the Substrate Mixture ( $t = 0$  s), by which time low levels of substrate hydrolysis were manifest. The inert *N*-acetylated hydrolysis products 27<sup>47</sup> and 5 or 24 were used as internal standard to quantify hydrolysis (Supporting Figures S2 and S4); SPE-MS assay results are a mean of independent triplicates ( $n = 3$ ; mean  $\pm$  SD).

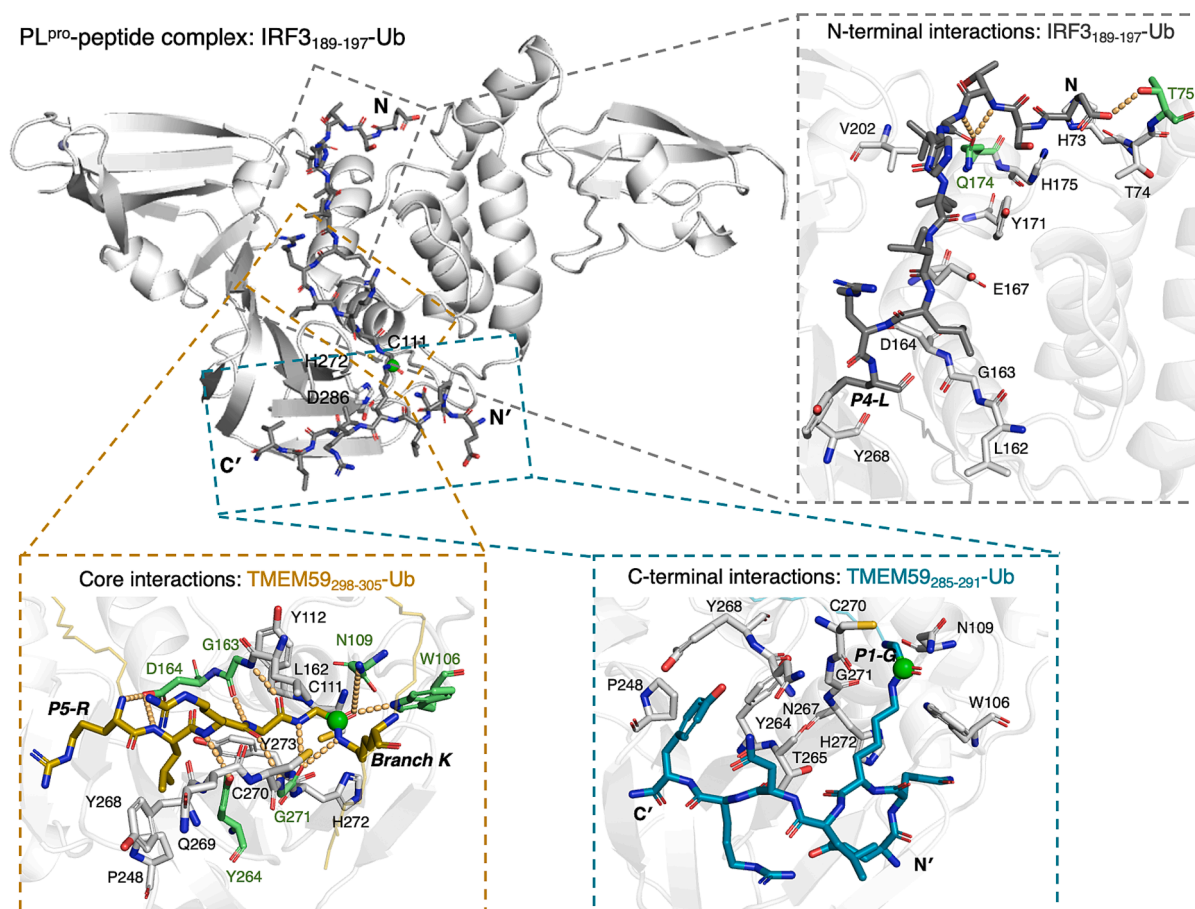
In all the modelled PL<sup>pro</sup>:substrate complexes, the residues C-terminal to the substrate LXGG motif manifest high conformational flexibility, with backbone RMSFs exceeding 4 Å (Supporting Figure S16). The V-shaped RMSF plots indicate that the peptide backbone is flexible at both its N- and C-termini (Supporting Figure S16). Nonetheless, transient hydrogen bonding and dispersion interactions were observed between the substrates and e.g. N109 of PL<sup>pro</sup> (Supporting Table S1), as well as with residues around Y268 which form a flexible  $\beta$ -hairpin loop (G266-G271), i.e. the blocking loop 2 (BL2), that controls substrate access to the active site.<sup>67-70</sup> In the apo form of PL<sup>pro</sup>, the BL2 loop is present in an open conformation; substrate binding likely triggers an induced fit mechanism resulting in closure of the BL2 loop, in a manner helping to orient the C-terminus of the UBL productively towards the active site.<sup>70</sup> The conformational changes in the BL2 loop associated with substrate binding to PL<sup>pro</sup> are conserved in, at least, some human DUBs, suggesting that both viral and human DUBs employ similar mechanisms to regulate substrate selectivity.<sup>71,72</sup> Note that SARS-CoV and SARS-CoV-2 PL<sup>pro</sup> inhibitors can also bind proximate to BL2 and alter or stabilize its conformation,<sup>39,67,70,73-76</sup> as preceded by the inhibitor-induced conformational changes in BL2 of human DUBs, such as USP7 and USP14.<sup>77</sup>

Our previous modelling studies have shown that the BL2 conformation is stable in its closed form in the PL<sup>pro</sup>:nsp2/3<sub>808-827</sub> (2) substrate complex, though alternating between open and closed forms with apo PL<sup>pro</sup>, or when the active site is occupied by a peptide that is not efficiently hydrolyzed.<sup>61</sup> We analyzed BL2 dynamics in all the nine modelled PL<sup>pro</sup>:substrate complexes by backbone RMSF of the most flexible residue in BL2 (i.e. Y268) (Fig. 7, Supporting Figures S25 and

S26) and the Y268-P248  $\alpha$ - $\alpha$  distance (Supporting Figures S27 and S28); in the latter case, a distance of  $>12$  Å was considered to be indicative of an open BL2 conformation. The results reveal that BL2 remains stably closed in the PL<sup>pro</sup>:IRF3<sub>189-197</sub>-ISG15 (4) and PL<sup>pro</sup>:IRF3<sub>189-197</sub>-Ub (15) complexes (Fig. 7), which are the most efficient substrates among the modelled peptides (Fig. 3). By contrast, BL2 is flexible in the modelled PL<sup>pro</sup> complex with the relatively inefficient substrate IRF3<sub>189-197</sub>-URM1 (17), regardless of the charge state of the P3 histidine imidazole group. Similarly, in the modelled PL<sup>pro</sup>:IRF3<sub>189-197</sub>-SUMO1 (18) complex, BL2 frequently opens and closes (Supporting Figure S28); note that 18 is not a PL<sup>pro</sup> substrate (Fig. 3). The latter observation is likely a result of the P4 leucine to glutamine substitution in 18; consistent with this proposal, previous reports have shown that the binding of hydrophobic sidechains to the PL<sup>pro</sup> S4 site induces BL2 closure.<sup>65,70,78,79</sup>

In the modelled PL<sup>pro</sup> complexes with the *N*<sup>ε</sup>-lysine-branched oligopeptides that mimic post-translational protein ubiquitylation (i.e. 15, 21, 22, and 23), BL2 remains rigidly closed for the relatively more efficient substrates (i.e. 15 and 21), as observed for the pp1a/lab-derived nsp2/3<sub>808-827</sub> peptide 2,<sup>61</sup> but not for the less efficient substrates (i.e. 22 and 23) (Fig. 7, Supporting Figures S27 and S28). These observations are in accord with the results on the stability of the complexes of PL<sup>pro</sup> and *N*<sup>ε</sup>-lysine-branched oligopeptides mimicking post-translational protein ISG15ylation. In particular, it appears that hydrophobic residues C-terminal to the P1' lysine residue can affect BL2 flexibility (Supporting Figures S13, S23, S24, and S29-S31), in accord with the experimentally observed preference of PL<sup>pro</sup> for hydrophobic residues C-terminal to the substrate P1' lysine residue involved in the





**Fig. 6. Modelled interactions of PL<sup>pro</sup> with lysine-branched oligopeptides.** Views of clustering-derived representative PL<sup>pro</sup>:oligopeptide complex structures exemplifying the overall binding mode of the oligopeptides and their interactions with PL<sup>pro</sup> N-terminal and C-terminal to the S4-S1 sites. Peptides are colored as in Fig. 3, with the P1 scissile amide carbons as green spheres; the N-terminus of the UBL-derived fragment and the N- and C-termini of the S' binding fragment are labelled N, N', and C', respectively. PL<sup>pro</sup> residues interacting via hydrogen bonding with the oligopeptide are in lime (the calculated occurrence is  $\geq 25\%$ ; orange dashes; not necessarily present in the frame). PL<sup>pro</sup> residues within 4 Å of the peptide residues that are calculated to contribute  $\geq 0.5$  kcal mol<sup>-1</sup> binding energy, are in white.

isopeptide amide bond (Fig. 4).

The combined modelling and experimental results reveal the importance of the conformational dynamics of the BL2 loop; it not only enables productive substrate binding in the active site, but also governs substrate selectivity. Importantly, the BL2 flexibility and opening tendency observed in the MD simulations correlate with the experimentally observed substrate selectivity, i.e. conformational BL2 flexibility in the PL<sup>pro</sup>:substrate complex appears to result in inefficient catalysis (Fig. 7). Although substrate peptide dissociation was not observed over the course of the MD simulations, the observations suggest that a sustained closed state of BL2 may be necessary for initiating productive catalysis, potentially because BL2 is proximal to H272 which is part of the catalytic triad.<sup>68</sup> Thus, the observation that the PL<sup>pro</sup>-catalyzed hydrolysis of the linear pp1a/1ab-derived oligopeptide nsp2/3<sub>808-827</sub> (2) and the N<sup>ε</sup>-lysine-branched oligopeptide substrates do not appear to affect each other when incubated in the same reaction vessel (Fig. 5), implies that inefficient substrate binding to PL<sup>pro</sup> results in dissociation being faster than substrate hydrolysis.

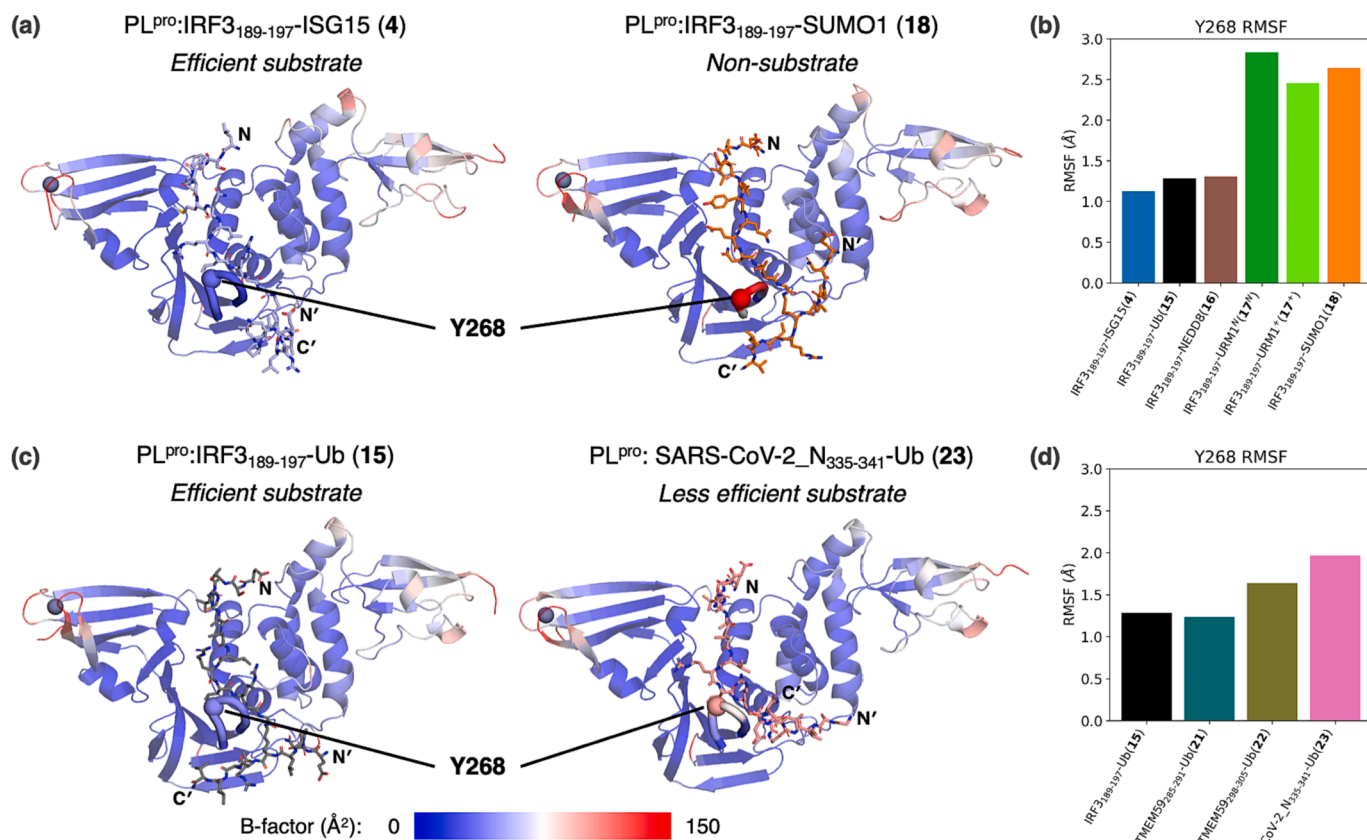
### 3. Conclusions

SARS-CoV-2 PL<sup>pro</sup> catalyzes the hydrolysis of peptide amide bonds that are C-terminal to three LXGG motifs in pp1a/1ab to release functional nsp1-3.<sup>1,2</sup> By contrast with M<sup>pro</sup>, PL<sup>pro</sup> is also a deubiquitinase (DUB) that catalyzed the hydrolysis of isopeptide amide bonds of human proteins that are post-translationally modified with Ub/UBLs, in

particular ISG15.<sup>39,40</sup> Thus, the DUB substrate preference of PL<sup>pro</sup> is apparently similar to that of the human ubiquitin specific peptidase 18 (USP18), a DUB which also employs a nucleophilic cysteine and which selectively catalyzes protein deISG15ylations.<sup>35,45</sup> It has been proposed that the DUB activity of PL<sup>pro</sup> modulates virulence via catalyzing the deISG15ylation of post-translationally ISG15ylated IRF3.<sup>39</sup> However, to our knowledge, the DUB activity of isolated recombinant PL<sup>pro</sup> has not yet been validated with oligopeptide substrates *in vitro*.

We developed SPE-MS assays that directly monitor the hydrolysis of the isopeptide amide bond of synthetic N<sup>ε</sup>-lysine-branched oligopeptides, which mimic post-translational protein modification by Ub/UBLs, catalyzed by isolated recombinant SARS-CoV-2 PL<sup>pro</sup>, extending the scope of our reported PL<sup>pro</sup> SPE-MS assays with linear oligopeptides as substrates to N<sup>ε</sup>-lysine-branched oligopeptide substrates.<sup>47</sup> The SPE-MS assays compare favorably to reported spectroscopic- or MS-based DUB assays, because these do not employ substrates that bind to the PL<sup>pro</sup> S' sites,<sup>39,40,48,49</sup> and are associated with comparatively high costs of the substrates.<sup>51</sup> Due to variations in the catalytic efficiency of different PL<sup>pro</sup> batches, we thus determined the rank order of substrate preference by quantifying product formation relative to internal standards.

Importantly, our results with N<sup>ε</sup>-lysine-branched oligopeptides support cellular studies concerning the DUB activities of PL<sup>pro</sup>, including its reported preference for catalyzing the hydrolysis of isopeptide amide bonds C-terminal to the LRGG motif of ISG15 rather than Ub/Ub<sub>2</sub> and, in particular, NEDD8.<sup>39,40</sup> The rank order observed for PL<sup>pro</sup>-catalyzed hydrolysis of K<sub>193</sub>-branched IRF3-derived oligopeptides mimicking post-



**Fig. 7.** BL2 flexibility in PL<sup>pro</sup>:oligopeptide complexes is indicative of poor hydrolysis efficiency. Views of clustering-derived representative structures of (a) the PL<sup>pro</sup>:IRF3<sub>189-197</sub>-ISG15 (4) and PL<sup>pro</sup>:IRF3<sub>189-197</sub>-SUMO1 (18) complexes and (c) the PL<sup>pro</sup>:IRF3<sub>189-197</sub>-Ub (15) and PL<sup>pro</sup>:SARS-CoV-2<sub>N</sub>-Ub (23) complexes. PL<sup>pro</sup> residues are colored based on MD-derived backbone B-factors; red represents high B-factors, which is indicative of relatively high residue flexibility. The bar plots show comparisons of Y268 backbone RMSF for PL<sup>pro</sup> in complex with (b) oligopeptides 4, 15, 16, 17, and 18, and (d) oligopeptides 15, 21, 22, and 23. Residues forming BL2 (G266-G271) are shown as a thickened loop, with Y268 Cα being shown as a sphere. The N-terminus of the UBL-derived fragment and the N- and C-termini of the S' binding fragment are labelled N, N', and C', respectively.

translational modification by UBLs was: IRF3<sub>189-197</sub>-ISG15 (4) > IRF3<sub>189-197</sub>-Ub (15) > IRF3<sub>189-197</sub>-NEDD8 (16) > IRF3<sub>189-197</sub>-URM1 (17) > IRF3<sub>189-197</sub>-SUMO1 (18), the latter of which was not a substrate of PL<sup>pro</sup> (Fig. 3a). The results thus imply that the substrate preference of PL<sup>pro</sup> not only depends on the fold of the substrate, interactions of the substrate at the PL<sup>pro</sup> S1-S4 sites,<sup>42,65</sup> and on allosteric interactions with PL<sup>pro</sup>,<sup>42-44</sup> but also on the sequence identity of the region of the substrate which binds to the active site beyond S1-S4. This proposal is further supported by the pronounced effect of amino acid substitutions in the N<sup>ε</sup>-lysine-branched oligopeptide substrates on PL<sup>pro</sup> catalysis (Fig. 4), as well as by reported work on the hydrolysis of linear oligopeptides catalyzed by SARS-CoV-2 PL<sup>pro</sup> and PL<sup>pro</sup>s of other coronaviruses.<sup>80,81</sup>

The combined results show that the primary sequence of both the peptide fragment derived from the C-terminus of the UBL and that of the post-translationally modified human protein affects the efficiency of PL<sup>pro</sup> catalysis, the latter potentially by binding to the S' sites of PL<sup>pro</sup>, possibly in a substrate context-dependent manner. Hence, SPE-MS assays with oligopeptides have potential to predict the efficacy with which PL<sup>pro</sup> catalyzes the hydrolysis of post-translationally modified (human) proteins, although factors other than the primary sequence also contribute to catalysis efficiency (see above). It will be useful to perform studies with the N<sup>ε</sup>-lysine-branched oligopeptides and isolated recombinant SARS-CoV PL<sup>pro</sup> and MERS-CoV PL<sup>pro</sup>, as well as reported variants of SARS-CoV-2 PL<sup>pro</sup>, to inform on different substrate preferences and on how these differences manifest in altered virulence. Similarly, the substrate preferences of viral proteases other than coronavirus PL<sup>pro</sup>s, that have been reported to catalyze protein deISG15ylation and/

or deubiquitylation,<sup>82-84</sup> should be examined.

The results indicate that PL<sup>pro</sup> may catalyze the deISG15ylation of IRF3 regardless of the ISG15ylation site (Fig. 2c); studies with SARS-CoV-2 infected human cells are required to investigate the significance of this observation in a cellular context. Note that the ISG15ylation of IRF3 at the N<sup>ε</sup>-amino group of K193, K366, and, by implication, K360 may compete with other post-translational modifications, e.g. ubiquitylation at K193,<sup>54</sup> and/or acetylation<sup>85</sup> and methylation<sup>86</sup> at K366, in a manner potentially affecting PL<sup>pro</sup> catalysis. SPE-MS assay results with N<sup>ε</sup>-lysine-branched oligopeptides mimicking the ISG15ylation of 4EHP, the ISG15ylation of which has been proposed to control translation during immune response,<sup>59</sup> suggest that ISG15ylated 4EHP may be a substrate of SARS-CoV-2 PL<sup>pro</sup> in cells (Fig. 3b). The relevance of the PL<sup>pro</sup>-catalyzed deISG15ylation of ISG15ylated 4EHP for virulence should thus also be probed in cellular studies.

It is likely that PL<sup>pro</sup> not only catalyzes the deISG15ylation of IRF3 and 4EHP (Fig. 3), but also of other human proteins.<sup>87</sup> Further substrate selectivity studies are, however, currently hampered by the lack of information on the exact ISG15ylation sites of human proteins.<sup>35,55-58</sup> Note that evidence suggests that residues other than lysine residues may be ISG15ylated, e.g. cysteine residues<sup>88,89</sup>; thus, the ability of SARS-CoV-2 PL<sup>pro</sup> to catalyze the deISG15ylation of other residues than lysine should be subject of future work. Considering that protein ubiquitylation is apparently more widespread than ISG15ylation, it is possible that human ubiquitylated proteins may be identified which are more efficient substrates than ISG15ylated IRF3 *in vitro* and in cells.

Viral proteins can also be ISG15ylated by host proteins, in a manner which may impede their function, e.g. during replication.<sup>90-92</sup> PL<sup>pro</sup> may

catalyze the deISG15ylation of ISG15ylated viral proteins to counteract the host innate immune response. Although, to our knowledge, little information is currently available on SARS-CoV-2 protein ISG15ylation, multiple ubiquitinylation sites have been identified in SARS-CoV-2 proteins in proteomic MS studies.<sup>62</sup> Our work provides MS evidence that isolated recombinant PL<sup>pro</sup> catalyzes the hydrolysis of the isopeptide amide bond of an N<sup>ε</sup>-lysine-branched oligopeptide based on a reported ubiquitinylation site of the SARS-CoV-2 N protein, *i.e.* SARS-CoV-2<sub>N335-341</sub>-Ub 23 (Fig. 3c). It is thus possible that PL<sup>pro</sup> catalyzes the deubiquitinylation of at least some of the (poly)ubiquitinated lysine residues of viral proteins to *e.g.* regulate protein function or to prevent protein degradation via the host proteasome, as (poly)ubiquitinylation can be a signal for proteasomal degradation and the removal of the (poly)ubiquitin may hence stabilize the viral protein; conversely, PL<sup>pro</sup> has potential to also catalyze the deISG15ylation of ISG15ylated viral proteins.

In general, it appears that viral proteases may have evolved to catalyze deISG15ylation reactions as a strategy to evade the host innate immune response, at least during early stages of infection.<sup>92–94</sup> Interestingly, viral proteases catalyze deISG15ylation reactions at different sites, *e.g.* the leader protease of foot-and-mouth-disease virus (FMDV) is reported to catalyze the hydrolysis of ISG15 N-terminal to diglycine of the LRGG motif,<sup>82</sup> opposed to SARS-CoV-2 PL<sup>pro</sup> which catalyzes the hydrolysis of ISG15 C-terminal to diglycine of the LRGG motif. Viral proteases have also been reported to catalyze the direct hydrolysis of the main chain of interferon regulatory factors,<sup>95–97</sup> including that of IRF3 by SARS-CoV-2 PL<sup>pro</sup>,<sup>18</sup> further highlighting the importance of modulating (post-translationally modified) interferon regulatory factors such as IRF3 during viral infections.

At least in most cases, the identity of the substrate residues binding to the S sites affects PL<sup>pro</sup> catalysis to a relatively greater extent than the identity of the substrate residues binding to the S' sites (Fig. 2). This observation may reflect the different substrate selectivities of PL<sup>pro</sup> and M<sup>pro</sup>. M<sup>pro</sup> apparently requires the presence of specific residues (*i.e.*, S/A/N) at the S1' site for efficient catalysis, though its substrate requirements for S2'-S4' appear to be less stringent.<sup>98</sup> Defining the full substrate scope of M<sup>pro</sup> is, however, the subject of ongoing investigations.<sup>99,100</sup> Notably, NEMO, the M<sup>pro</sup>-catalyzed hydrolysis of which has been reported to induce the death of brain endothelial cells,<sup>16,17</sup> has a valine residue at the P1' position, indicating that the presence of (L/F/V)Q(S/A/N) motifs in human proteins may not be sufficient to predict the efficiency of M<sup>pro</sup> catalysis.

Eleven sites in pp1a/1ab fulfil the apparently somewhat narrower substrate requirements for M<sup>pro</sup>, whereas only three sites in pp1a/1ab fulfil the substrate requirements for the apparently more promiscuous PL<sup>pro</sup>.<sup>1,2</sup> Human proteases that accept (L/F/V)Q(S/A/N) motifs as substrates are currently unknown, suggesting that M<sup>pro</sup> may have evolved to exclusively catalyze the release of those nsps from pp1a/1ab that are directly involved in replication and transcription (*e.g.* nsp12, which has a RNA-dependent RNA polymerase domain, and the nsp13 helicase), under strict spatiotemporal control. By contrast, in principle, human DUBs could catalyze the hydrolysis of pp1a/1ab at three sites, releasing *inter alia* nsp1 and nsp2 which have been proposed to help the virus evade the host immune system,<sup>101</sup> and thus compete with PL<sup>pro</sup> for viral substrates. The apparent lower substrate specificity of PL<sup>pro</sup> may reflect its functions in enabling the virus to counteract the host immune system, with respect to catalyzing both the release of nsp1 and nsp2 and the deISG15ylation of post-translationally modified human and, potentially, viral proteins, a process which might not require strict spatiotemporal control and thus tolerates reduced substrate specificity.

The ability of PL<sup>pro</sup>, but not M<sup>pro</sup> (as far as is known), to catalyze hydrolysis of isopeptide as well as peptide amide bonds is striking. The comparison of substrate and non-substrate binding modes for the two proteases is thus of interest, because the precise nature of the dynamic interactions that regulate the different substrate selectivities of PL<sup>pro</sup> and M<sup>pro</sup> is not defined. Reported PL<sup>pro</sup> structures reveal that substrate

binding can induce conformational changes of flexible regions proximate to the active site, including the PL<sup>pro</sup> blocking loop 2 (BL2) which is a key factor in regulating the accessibility of potential substrates to the active site.<sup>70</sup> The role of the PL<sup>pro</sup> BL2 in regulating substrate recognition is preceded by BL2 in, at least some, human DUBs, including *e.g.* USP7 and USP14;<sup>71,72</sup> however, in human DUBs, post-translational modification of BL2 residues may further alter catalysis.<sup>102</sup> It appears that the PL<sup>pro</sup> substrate selectivity is not only a result of the affinity of the substrate to bind to S sites and, to a lesser extent, S' sites, but also of the ability of a potential substrate to stabilize the closed conformation of BL2, so enabling a catalytically productive enzyme-substrate conformation which may be achieved via interactions with PL<sup>pro</sup> that may be relatively remote from the active site, including by specific interactions of hydrophobic substrate residues with the S' sites (Fig. 6).

Movement of active site bordering loops to bind, enclose, and/or orientate substrates (and reject non-substrates) are common in enzyme catalysis. For example, substrate binding can induce conformational changes in flexible regions of M<sup>pro</sup> that directly contribute to shaping the S sites, in particular of an  $\alpha$ -helix (T45-L50), which contributes to forming the S2 site, and of a loop (D187-Q192), which contributes to forming the S4 site.<sup>98,103–106</sup> The presence of a BL2-type loop in PL<sup>pro</sup>, but not M<sup>pro</sup>, is interesting. Whether or not the presence of this loop affects the ability of PL<sup>pro</sup> to accept different types of substrates (peptide and isopeptide amide bonds), whilst manifesting an apparently narrower selectivity with respect to polyprotein pp1a/1ab hydrolysis, is unclear. In this regard, it will be of interest to investigate how the substrate selectivities of PL<sup>pro</sup> homologues have changed during the course of viral evolution, both with respect to pp1a/1ab and proteins in different host organisms, *e.g.* SARS-CoV-2 PL<sup>pro</sup> prefers ISG15 as a substrate, whereas SARS-CoV PL<sup>pro</sup> prefers diubiquitin.<sup>39,40</sup> Notably, the efficiency and selectivity of MERS-CoV PL<sup>pro</sup> catalysis has been proposed to be distinct from that of SARS-CoV and SARS-CoV-2 PL<sup>pro</sup>, in part as a result of amino acid variations in the BL2 loop.<sup>48,74</sup>

SPE-MS assays have been used to investigate post-translational protein modifications and catalysis of nucleophilic cysteine enzymes, including isolated recombinant SARS-CoV-2 PL<sup>pro</sup> and M<sup>pro</sup>.<sup>10,47,49,107–113</sup> Our SPE-MS assays employing synthetic N<sup>ε</sup>-lysine-branched oligopeptides as substrates of isolated recombinant PL<sup>pro</sup> have enabled studies on the substrate scope and substrate selectivity of PL<sup>pro</sup> and thus helped to characterize the DUB activity of PL<sup>pro</sup>. The combination of analogous MS-based assays with synthetic N<sup>ε</sup>-lysine-branched oligopeptides as substrates of isolated recombinant human DUBs, together with proteomic studies, will enable studies on the substrate preferences of human DUBs, such as USP18 which also catalyzes protein deISG15ylation, and which is a current medicinal chemistry target.<sup>35,45</sup>

## 4. Experimental section

### 4.1. Production and purification of isolated recombinant SARS-CoV-2 PL<sup>pro</sup>

The PL<sup>pro</sup> domain of the SARS-CoV-2 nsp3 (E746-T1063) was produced using *E. coli* Lemo21(DE3) cells and purified as reported previously.<sup>47</sup>

### 4.2. Peptide synthesis

Linear oligopeptides were prepared by solid phase peptide synthesis (SPPS) using a Liberty Blue peptide synthesizer (CEM Microwave Technology Ltd.), as reported for the synthesis of the oligopeptides 1–3.<sup>47</sup>

The N<sup>ε</sup>-lysine-branched oligopeptides mimicking post-translationally modified proteins were synthesized by microwave-assisted SPPS using the Fmoc-protection strategy from the C- to N-terminus on Rink Amide MBHA resin (AGTC Bioproducts Ltd.; loading: 0.6–0.8 mmol/g) similar to reported procedures.<sup>47</sup> Initially, a linear



oligopeptide was synthesized; *N*-Fmoc lysine with a 4-methyltrityl (Mtt) protected *N*<sup>ε</sup>-amine was used at the site of the lysine branching, and an amino acid with a Boc-protected *N*<sup>α</sup>-amine was used at the N-terminus. Branching was introduced via selective deprotection of the lysine *N*<sup>ε</sup>-Mtt group of the resin-bound peptides using 1%<sub>v/v</sub> trifluoroacetic acid (TFA) and 2%<sub>v/v</sub> triisopropylsilane (TIPS) in dichloromethane at ambient temperature while shaking (300 rpm) (5 cycles; after each cycle, the deprotection mixture was removed and the resin was washed with dichloromethane). Following Mtt deprotection, the UBL-derived peptide fragment was synthesized by SPPS from the C- to N-terminus starting with the free *N*<sup>ε</sup>-amino group of the resin-bound peptides at the C-terminus.

After completion of the synthesis, the resin-bound peptides were washed with dichloromethane and subsequently cleaved from the resin and simultaneously deprotected using a mixture of trifluoroacetic acid, triisopropylsilane, 1,3-dimethoxybenzene, and water (92.5/2.5/2.5/2.5%<sub>v/v</sub>, respectively). Solids were separated; the remaining clear solution was diluted with diethyl ether (45 mL/0.1 mmol resin). After incubation for 30 min at 0 °C, the mixture was centrifuged for 10 min using a Beckman Coulter Allegra X-30R centrifuge equipped with a SX4400 rotor (4500 rpm); the supernatant was discarded. The solid residue was dissolved in a water/acetonitrile mixture, frozen using liquid N<sub>2</sub>, and then lyophilized. The dried crude product was dissolved in a water/acetonitrile mixture, filtered, and purified using a semi-preparative HPLC machine (Shimadzu UK Ltd.) equipped with a reverse phase column (Gemini 00G-4454-UO-AX; phase: NX-C18). A linear gradient (typically 2–47%<sub>v/v</sub> over 38 min) of acetonitrile in milli-Q grade water (each containing 0.1%<sub>v/v</sub> trifluoroacetic acid) was used as eluent. Fractions were analyzed by SPE-MS and those containing the pure peptide were combined and lyophilized. Sequences, mass spectra, and purification characteristics of the synthetic oligopeptides are shown in Supporting Figure S1.

#### 4.3. SPE-MS assays

PL<sup>pro</sup> assays for turnover and competition experiments were performed in 96-well polypropylene assay plates (Greiner), either with a 1.0 or 0.5 mL final reaction volume, using isolated recombinant SARS-CoV-2 PL<sup>pro</sup> (0.2 μM), substrate peptide(s) (2.0 μM), inert *N*-acetylated standard peptides (0.2 μM, as indicated in the individual experiments) in buffer (50 mM Tris, pH 8.0) at ambient temperature. PL<sup>pro</sup> catalysis was directly monitored using SPE-MS.<sup>47</sup> The RapidFire RF 365 high-throughput sampling robot used was programmed to aspirate samples from the reaction mixture at the time intervals indicated in the individual experiments.

MS-analyses were performed using a RapidFire RF 365 high-throughput sampling robot (Agilent) attached to an iFunnel Agilent 6550 accurate mass quadrupole time-of-flight (Q-TOF) mass spectrometer operated in the positive ionization mode.<sup>47</sup> Assay samples were aspirated under vacuum for 0.6 s and loaded onto a C4 solid phase extraction (SPE) cartridge. After loading, the C4 SPE cartridge was washed with 0.1%<sub>v/v</sub> aqueous formic acid to remove non-volatile buffer salts (5.5 s, 1.5 mL/min). The peptide was eluted from the SPE cartridge with 0.1%<sub>v/v</sub> aqueous formic acid in 85/15%<sub>v/v</sub> acetonitrile/water into the mass spectrometer (5.5 s, 1.25 mL/min) and the SPE cartridge re-equilibrated with 0.1%<sub>v/v</sub> aqueous formic acid (0.5 s, 1.25 mL/min). The mass spectrometer parameters were: capillary voltage: 4000 V; nozzle voltage: 1000 V; fragmentor voltage: 365 V; gas temperature: 280 °C; gas flow: 13 L/min; sheath gas temperature: 350 °C; sheath gas flow: 12 L/min.

For data analysis and to quantify product formation, the charge states of both the C-terminal and N-terminal product peptides and the corresponding *N*-acetylated C-terminal and N-terminal product peptides (i.e. **5**, **12**, and **24**), which were used as internal standards, were used to extract ion chromatogram data (*m/z* +1 for **12** and the corresponding product peptide; *m/z* +2 for **5** and **24**, as well as for the corresponding

product peptides); peak areas were integrated using the RapidFire Integrator software (Agilent). Data were exported into Microsoft Excel and used to calculate the product peptide concentrations using the equation: peptide concentration = 0.2 μM × (integral C- or N-terminal product peptide) / (integral *N*-acetylated C- or N-terminal product peptide).

#### 4.4. Preparation of PL<sup>pro</sup>:oligopeptide models

A reported SARS-CoV-2 PL<sup>pro</sup> structure (PDB ID: 6WX4<sup>65</sup>) was prepared for modelling as described (Supporting Table S2).<sup>61</sup> Based on reported quantum mechanics/molecular mechanics-umbrella sampling (QM/MM-US) calculations on proton transfer processes in the PL<sup>pro</sup> catalytic triad,<sup>61</sup> C111 was modelled in its deprotonated form, H272 as doubly protonated, and D286 as deprotonated.

The *N*<sup>ε</sup>-lysine-branched oligopeptides complexed with PL<sup>pro</sup> were constructed using reported AutoDock CrankPep (ADCP)<sup>66</sup>-docked conformations of the linear oligopeptide nsp2/3<sub>808-827</sub> (**2**).<sup>61</sup> The residues N-terminal to the scissile amide in the top ranked pose (d2\_01, reported nomenclature<sup>61</sup>), which was successful in placing P4-P1 LKGG in the respective S4-S1 subsites, were used to build the N-terminal ubiquitin (Ub) or Ub-like modifier (UBL) derived fragment of the *N*<sup>ε</sup>-lysine-branched oligopeptides. The C-terminal fragment of the *N*<sup>ε</sup>-lysine-branched oligopeptides was built based on the 48th-ranked pose of **2** (d2\_48), which did not pass through S4-S1, but in which the P4' lysine sidechain *N*<sup>ε</sup> amine was proximate to C111 (Supporting Figure S5). The lysine side chain branching was modelled by linking the C<sup>ε</sup> atom of the d2\_48 P4' lysine sidechain to the carboxamide N atom of the d2\_01 P1' alanine. To prepare each of the eight branched peptides in IRF3<sub>189-197</sub>-(ISG15/Ub/NEDD8/URM1/SUMO1) and (TMEM59<sub>285-291</sub>/TMEM59<sub>298-305</sub>/SARS-CoV-2<sub>N335-341</sub>)-Ub (Fig. 3), the two poses of **2** were modified to the lengths of the N- and C-terminal fragments. The residues were modified to the target sequences using the mutagenesis tool of PyMOL (open source, v. 2.3.0),<sup>114</sup> selecting the least sterically clashing backbone-dependent rotamer in each case.<sup>115</sup> The N-termini of both fragments were uncapped, while the C-terminus of the IRF3<sub>189-197</sub>/TMEM59<sub>285-291</sub>/TMEM59<sub>298-305</sub>/SARS-CoV-2<sub>N335-341</sub> fragment was NH<sub>2</sub>-capped. Peptide histidine residues were modelled in their neutral state, with *N*<sup>ε</sup> protonated (HIE in AMBER nomenclature),<sup>116</sup> except for IRF3<sub>189-197</sub>-URM1 where both the neutral HIE and the positively charged, doubly protonated (HIP) states of the P3 histidine residue were considered, denoted as URM1<sup>N</sup> and URM1<sup>+</sup> respectively.

#### 4.5. Parametrization and molecular dynamics (MD)

MD simulations of the PL<sup>pro</sup>:oligopeptide complexes were performed in GROMACS (v. 2019.2, 2020.4),<sup>117</sup> using the AMBERFF99SB-ILDN/TIP3P forcefield.<sup>116,118</sup> Non-bonded Zn parameters were employed.<sup>119</sup> For the non-standard *N*<sup>ε</sup>-branched lysine residue, *N*-butyl acetamide (Supporting Figure S32) was used as a model for parametrization with AMBER atom types and RESP charges (Gaussian16 and antechamber in Amber18).<sup>120–124</sup> The new residue, designated as “LYC”, was incorporated into GROMACS (Supporting Figure S33), with its *N*<sup>ε</sup> atom bonded to the P1 Gly C atom. Amide improper dihedral parameters centered at the GLY\_C and LYC\_NZ atoms were included to maintain amide group planarity.

Each of the nine PL<sup>pro</sup>:oligopeptide complexes (eight unique sequences, with URM1<sup>N</sup> and URM1<sup>+</sup> for two possible charge states of the P3 histidine group in **17**) was centered in a rhombic dodecahedral box with at least 1.0 nm separation from box edges, solvated, neutralized with sodium/chloride ions (110,788–110,847 atoms in total), and minimized until the maximum force was below 1000 kJ mol<sup>−1</sup> nm<sup>−1</sup>. From the minimized system, three replicas were initiated using random velocities at 298.15 K, subjected to 200 ps (1 fs step) restrained NVT equilibration at 298.15 K, followed by 200 ps (1 fs step) NPT equilibration at 298.15 K and 1.0 bar. The equilibrated complexes were



subjected to 200 ns production MD (2 fs step). A velocity-rescaling thermostat with a stochastic term (time constant 0.1 ps; protein and non-protein coupled separately)<sup>125</sup> and a Parrinello-Rahman barostat (time constant 2 ps) were used.<sup>126,127</sup> Long-range electrostatic interactions were calculated by smooth Particle-mesh Ewald (1 nm cut-off).<sup>128,129</sup> Van der Waals interactions were cut off at 1 nm.

The 3 × 200 ns MD trajectories were fitted based on the PL<sup>pro</sup> backbone and analyzed using GROMACS tools (v 2019.2).<sup>117</sup> A hydrogen bond was defined on observation of a donor-acceptor distance <3.5 Å and hydrogen-donor-acceptor angle <30°. To obtain representative structures, clustering was performed with a 3 Å RMSD cut-off of the peptide backbone, using the gromos algorithm.<sup>130</sup> Per residue decomposition of PL<sup>pro</sup>:oligopeptide binding energies calculated by molecular mechanics/generalized Born surface area (MM/GBSA)<sup>131–134</sup> was performed on frames in 5 ns intervals, using MMPBSA.py (AMBER18) with an ionic strength of 0.15 M, mbondi2 radii, and igb = 5.<sup>124,135,136</sup>

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

Data will be made available on request.

## Acknowledgements

The investigators acknowledge the philanthropic support of the donors to the University of Oxford's COVID-19 Research Response Fund and King Abdulaziz University, Saudi Arabia, for funding. This research was funded in part by the Wellcome Trust (106244/Z/14/Z). For the purpose of open access, the author has applied a CC BY public copyright license to any Author Accepted Manuscript version arising from this submission. We thank Cancer Research UK (C8717/A18245) and the Biotechnology and Biological Sciences Research Council (BB/J003018/1 and BB/R000344/1) for funding. We thank Dr. J. J. A. G. Kamps for providing PL<sup>pro</sup> for preliminary studies. H.T.H.C. thanks the Clarendon Fund, New College Oxford, and the EPSRC Centre for Doctoral Training in Synthesis for Biology and Medicine (EP/L015838/1) for a studentship, generously supported by AstraZeneca, Diamond Light Source, Defence Science and Technology Laboratory, Evotec, GlaxoSmithKline, Janssen, Novartis, Pfizer, Syngenta, Takeda, UCB and Vertex. This project made use of computer time on HPC granted via the UK High-End Computing Consortium for Biomolecular Simulation, HECBioSim (<http://hecbiosim.ac.uk>), supported by EPSRC (grant no. EP/R029407/1).

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.bmc.2023.117498>.

## References

- V'kovski P, Kratzel A, Steiner S, Stalder H, Thiel V. Coronavirus biology and replication: implications for SARS-CoV-2. *Nat Rev Microbiol*. 2021;19:155–170.
- Thiel V, Ivanov KA, Putics Á, et al. Mechanisms and enzymes involved in SARS coronavirus genome expression. *J Gen Virol*. 2003;84:2305–2315.
- Li G, Hilgenfeld R, Whitley R, De Clercq E. Therapeutic strategies for COVID-19: progress and lessons learned. *Nat Rev Drug Discov*. 2023;22:449–475.
- Cannalire R, Cerchia C, Beccari AR, Di Leva FS, Summa V. Targeting SARS-CoV-2 proteases and polymerase for COVID-19 treatment: state of the art and future opportunities. *J Med Chem*. 2022;65:2716–2746.
- Ghosh AK, Mishevich JL, Mesecar A, Mitsuya H. Recent drug development and medicinal chemistry approaches for the treatment of SARS-CoV-2 infection and COVID-19. *ChemMedChem*. 2022;17:e202200440.
- Lv Z, Cano KE, Jia L, Drag M, Huang TT, Olsen SK. Targeting SARS-CoV-2 proteases for COVID-19 antiviral development. *Front Chem*. 2022;9, 819165.
- Ton A-T, Pandey M, Smith JR, Ban F, Fernandez M, Cherkasov A. Targeting SARS-CoV-2 papain-like protease in the postvaccine era. *Trends Pharmacol Sci*. 2022;43: 906–919.
- Tan H, Hu Y, Jadhav P, Tan B, Wang J. Progress and challenges in targeting the SARS-CoV-2 papain-like protease. *J Med Chem*. 2022;65:7561–7580.
- Owen DR, Allerton CMN, Anderson AS, et al. An oral SARS-CoV-2 Mpro inhibitor clinical candidate for the treatment of COVID-19. *Science*. 2021;374:1586–1593.
- Unoh Y, Uehara S, Nakahara K, et al. Discovery of S-217622, a noncovalent oral SARS-CoV-2 3CL protease inhibitor clinical candidate for treating COVID-19. *J Med Chem*. 2022;65:6499–6512.
- Mielech AM, Kilianski A, Baez-Santos YM, Mesecar AD, Baker SC. MERS-CoV papain-like protease has deISGylating and deubiquitinating activities. *Virology*. 2014;450–451:64–70.
- Yang X, Chen X, Bian G, et al. Proteolytic processing, deubiquitinase and interferon antagonist activities of middle east respiratory syndrome coronavirus papain-like protease. *J Gen Virol*. 2014;95:614–626.
- Lindner HA, Lytynen V, Qi H, Lachance P, Ziomek E, Ménard R. Selectivity in ISG15 and ubiquitin recognition by the SARS coronavirus papain-like protease. *Arch Biochem Biophys*. 2007;466:8–14.
- Ratia K, Kilianski A, Baez-Santos YM, Baker SC, Mesecar A. Structural basis for the ubiquitin-linkage specificity and deISGylating activity of SARS-CoV papain-like protease. *PLoS Pathog*. 2014;10:e1004113.
- Chen Z, Wang Y, Ratia K, Mesecar AD, Wilkinson KD, Baker SC. Proteolytic processing and deubiquitinating activity of papain-like proteases of human coronavirus NL63. *J Virol*. 2007;81:6007–6018.
- Wenzel J, Lampe J, Müller-Fielitz H, et al. The SARS-CoV-2 main protease Mpro causes microvascular brain pathology by cleaving NEMO in brain endothelial cells. *Nat Neurosci*. 2021;24:1522–1533.
- Hameedi MA, Prates ET, Garvin MR, et al. Structural and functional characterization of NEMO cleavage by SARS-CoV-2 3CLpro. *Nat Commun*. 2022;13: 5285.
- Moustaqil M, Ollivier E, Chiu H-P, et al. SARS-CoV-2 proteases PLpro and 3CLpro cleave IRF3 and critical modulators of inflammatory pathways (NLRP12 and TAB1): implications for disease presentation across species. *Emerg Microbes Infect*. 2021;10:178–195.
- Reynolds ND, Aceves NM, Liu JL, et al. The SARS-CoV-2 SSHHPS recognized by the papain-like protease. *ACS Infect Dis*. 2021;7:1483–1502.
- Mohamud Y, Xue YC, Liu H, et al. The papain-like protease of coronaviruses cleaves ULK1 to disrupt host autophagy. *Biochem Biophys Res Commun*. 2021;540:75–82.
- Meyer B, Chiaravalli J, Gellenoncourt S, et al. Characterising proteolysis during SARS-CoV-2 infection identifies viral cleavage sites and cellular targets with therapeutic potential. *Nat Commun*. 2021;12:5553.
- Haas AL, Ahrens P, Bright PM, Ankel H. Interferon induces a 15-kilodalton protein exhibiting marked homology to ubiquitin. *J Biol Chem*. 1987;262:11315–11323.
- Loeb KR, Haas AL. The interferon-inducible 15-kDa ubiquitin homolog conjugates to intracellular proteins. *J Biol Chem*. 1992;267:7806–7813.
- Narasimhan J, Wang M, Fu Z, Klein JM, Haas AL, Kim J-J-P. Crystal structure of the interferon-induced ubiquitin-like protein ISG15. *J Biol Chem*. 2005;280: 27356–27365.
- Kang JA, Kim YJ, Jeon YJ. The diverse repertoire of ISG15: more intricate than initially thought. *Exp Mol Med*. 2022;54:1779–1792.
- Kumar S, Tomooka Y, Noda M. Identification of a set of genes with developmentally down-regulated expression in the mouse brain. *Biochem Biophys Res Commun*. 1992; 185:1155–1161.
- Kamitani T, Kito K, Nguyen HP, Yeh ETH. Characterization of NEDD8, a developmentally down-regulated ubiquitin-like protein. *J Biol Chem*. 1997;272: 28557–28562.
- Furukawa K, Mizushima N, Noda T, Ohsumi Y. A protein conjugation system in yeast with homology to biosynthetic enzyme reaction of prokaryotes. *J Biol Chem*. 2000;275:7462–7465.
- Van der Veen AG, Schorpp K, Schlieker C, et al. Role of the ubiquitin-like protein Urm1 as a noncanonical lysine-directed protein modifier. *Proc Natl Acad Sci USA*. 2011;108:1763–1770.
- van der Veen AG, Ploegh HL. Ubiquitin-like proteins. *Ann Rev Biochem*. 2012;81: 323–357.
- Hochstrasser M. Origin and function of ubiquitin-like proteins. *Nature*. 2009;458: 422–429.
- Dikic I, Schulman BA. An expanded lexicon for the ubiquitin code. *Nat Rev Mol Cell Biol*. 2023;24:273–287.
- Glickman MH, Ciechanover A. The ubiquitin-proteasome proteolytic pathway: destruction for the sake of construction. *Physiol Rev*. 2002;82:373–428.
- Lecker SH, Goldberg AL, Mitch WE. Protein degradation by the ubiquitin-proteasome pathway in normal and disease states. *J Am Soc Nephrol*. 2006;17:1807–1819.
- Jiménez Fernández D, Hess S, Knobloch K-P. Strategies to target ISG15 and USP18 toward therapeutic applications. *Front Chem*. 2020;7:923.
- Cappadocia L, Lima CD. Ubiquitin-like protein conjugation: structures, chemistry, and mechanism. *Chem Rev*. 2018;118:889–918.
- Lange SM, Armstrong LA, Kulathu Y. Deubiquitinases: from mechanisms to their inhibition by small molecules. *Mol Cell*. 2022;82:15–29.
- Komander D, Clague MJ, Urbé S. Breaking the chains: structure and function of the deubiquitinases. *Nat Rev Mol Cell Biol*. 2009;10:550–563.
- Shin D, Mukherjee R, Grewe D, et al. Papain-like protease regulates SARS-CoV-2 viral spread and innate immunity. *Nature*. 2020;587:657–662.

- 40 Klemm T, Ebert G, Calleja DJ, et al. Mechanism and inhibition of the papain-like protease, PLpro, of SARS-CoV-2. *EMBO J.* 2020;39, e106275.
- 41 Lei J, Kusov Y, Hilgenfeld R. Nsp3 of coronaviruses: structures and functions of a large multi-domain protein. *Antiviral Res.* 2018;149:58–74.
- 42 Patchett S, Lv Z, Rut W, et al. A molecular sensor determines the ubiquitin substrate specificity of SARS-CoV-2 papain-like protease. *Cell Rep.* 2021;36, 109754.
- 43 van Vliet VJE, Huynh N, Palà J, et al. Ubiquitin variants potently inhibit SARS-CoV-2 PLpro and viral replication via a novel site distal to the protease active site. *PLoS Pathog.* 2022;18:e1011065.
- 44 Wydorski PM, Osipiuk J, Lanham BT, et al. Dual domain recognition determines SARS-CoV-2 PLpro selectivity for human ISG15 and K48-linked di-ubiquitin. *Nat Commun.* 2023;14:2366.
- 45 Malakhov MP, Malakhova OA, Kim KI, Ritchie KJ, Zhang D-E. UBP43 (USP18) specifically removes ISG15 from conjugated proteins. *J Biol Chem.* 2002;277: 9976–9981.
- 46 Gold IM, Reis N, Glaser F, Glickman MH. Coronaviral PLpro proteases and the immunomodulatory roles of conjugated versus free Interferon Stimulated Gene product-15 (ISG15). *Semin Cell Dev Biol.* 2022;132:16–26.
- 47 Brewitz L, Kamps JJAG, Lukacik P, et al. Mass spectrometric assays reveal discrepancies in inhibition profiles for the SARS-CoV-2 papain-like protease. *ChemMedChem.* 2022;17:e202200016.
- 48 Freitas BT, Durie IA, Murray J, et al. Characterization and noncovalent inhibition of the deubiquitinase and deISGylase activity of SARS-CoV-2 papain-like protease. *ACS Infect Dis.* 2020;6:2099–2109.
- 49 Redhead MA, Owen CD, Brewitz L, et al. Bispecific repurposed medicines targeting the viral and immunological arms of COVID-19. *Sci Rep.* 2021;11:13208.
- 50 Ritorto MS, Ewan R, Perez-Oliva AB, et al. Screening of DUB activity and specificity by MALDI-TOF mass spectrometry. *Nat Commun.* 2014;5:4763.
- 51 Armstrong LA, Lange SM, Dee Cesare V, et al. Biochemical characterization of protease activity of Nsp3 from SARS-CoV-2 and its inhibition by nanobodies. *PLoS One.* 2021;16:e0253364.
- 52 Wu F, Zhao S, Yu B, et al. A new coronavirus associated with human respiratory disease in China. *Nature.* 2020;579:265–269.
- 53 Shi H-X, Yang K, Liu X, et al. Positive regulation of interferon regulatory factor 3 activation by Herc5 via ISG15 modification. *Mol Cell Biol.* 2010;30:2424–2436.
- 54 Chattopadhyay S, Kuzmanovic T, Zhang Y, Wetzel JL, Sen GC. Ubiquitination of the transcription factor IRF-3 activates RIPA, the apoptotic pathway that protects mice from viral pathogenesis. *Immunity.* 2016;44:1151–1161.
- 55 Giannakopoulos NV, Luo J-K, Papov V, et al. Proteomic identification of proteins conjugated to ISG15 in mouse and human cells. *Biochem Biophys Res Commun.* 2005;336:496–506.
- 56 Li C, Nelson TG, Vertegaal ACO, Thibault P. Proteomic strategies for characterizing ubiquitin-like modifications. *Nat Rev Methods Primers.* 2021;1:53.
- 57 Zhao C, Denison C, Huibregtse JM, Gygi S, Krug RM. Human ISG15 conjugation targets both IFN-induced and constitutively expressed proteins functioning in diverse cellular pathways. *Proc Natl Acad Sci USA.* 2005;102:10200–10205.
- 58 Zhang Y, Thery F, Wu NC, et al. The in vivo ISGylome links ISG15 to metabolic pathways and autophagy upon *Listeria monocytogenes* infection. *Nat Commun.* 2019;10:5383.
- 59 Okumura F, Zou W, Zhang D-E. ISG15 modification of the eIF4E cognate 4EHP enhances cap structure-binding activity of 4EHP. *Genes Dev.* 2007;21:255–260.
- 60 Glaser W, Skern T. Extremely efficient cleavage of eIF4G by picornaviral proteinases L and 2A in vitro. *FEBS Lett.* 2000;480:151–155.
- 61 Chan HTH, Brewitz L, Lukacik P, et al. Studies on the selectivity of the SARS-CoV-2 papain-like protease reveal the importance of the P2' proline of the viral polyprotein. *bioRxiv.* 2023. <https://doi.org/10.1101/2023.07.11.548309>.
- 62 Stukalov A, Girault V, Grass V, et al. Multilevel proteomics reveals host perturbations by SARS-CoV-2 and SARS-CoV. *Nature.* 2021;594:246–252.
- 63 Kuo C-J, Chao T-L, Kao H-C, et al. Kinetic characterization and inhibitor screening for the proteases leading to identification of drugs against SARS-CoV-2. *Antimicrob Agents Chemother.* 2021;65:e02577–02520.
- 64 Nakashima Y, Brewitz L, Tumber A, Salah E, Schofield CJ. 2-Oxoglutarate derivatives can selectively enhance or inhibit the activity of human oxygenases. *Nat Commun.* 2021;12:6478.
- 65 Rut W, Lv Z, Zmudzinski M, et al. Activity profiling and crystal structures of inhibitor-bound SARS-CoV-2 papain-like protease: A framework for anti-COVID-19 drug design. *Sci Adv.* 2020;6:eabd4596.
- 66 Zhang Y, Sanner MF. AutoDock CrankPep: combining folding and docking to predict protein-peptide complexes. *Bioinformatics.* 2019;35:5121–5127.
- 67 Báez-Santos YM, St. John SE, Mesecar AD. The SARS-coronavirus papain-like protease: structure, function and inhibition by designed antiviral compounds. *Antivir Res.* 2015;115:21–38.
- 68 Henderson JA, Verma N, Harris RC, Liu R, Shen J. Assessment of proton-coupled conformational dynamics of SARS and MERS coronavirus papain-like proteases: Implication for designing broad-spectrum antiviral inhibitors. *J Chem Phys.* 2020; 153, 115101.
- 69 Gao X, Qin B, Chen P, et al. Crystal structure of SARS-CoV-2 papain-like protease. *Acta Pharm Sin B.* 2021;11:237–245.
- 70 Fu Z, Huang B, Tang J, et al. The complex structure of GRL0617 and SARS-CoV-2 PLpro reveals a hot spot for antiviral drug discovery. *Nat Commun.* 2021;12:488.
- 71 Hu M, Li P, Li M, et al. Crystal structure of a UBP-family deubiquitinating enzyme in isolation and in complex with ubiquitin aldehyde. *Cell.* 2002;111:1041–1054.
- 72 Hu M, Li P, Song L, et al. Structure and mechanisms of the proteasome-associated deubiquitinating enzyme USP14. *EMBO J.* 2005;24:3747–3756.
- 73 Ratia K, Pegan S, Takayama J, et al. A noncovalent class of papain-like protease/deubiquitinase inhibitors blocks SARS virus replication. *Proc Natl Acad Sci USA.* 2008;105:16119–16124.
- 74 Lee H, Lei H, Santarsiero BD, et al. Inhibitor recognition specificity of MERS-CoV papain-like protease may differ from that of SARS-CoV. *ACS Chem Biol.* 2015;10: 1456–1465.
- 75 Osipiuk J, Azizi S-A, Dvorkin S, et al. Structure of papain-like protease from SARS-CoV-2 and its complexes with non-covalent inhibitors. *Nat Commun.* 2021;12:743.
- 76 Shen Z, Ratia K, Cooper L, et al. Design of SARS-CoV-2 PLpro inhibitors for COVID-19 antiviral therapy leveraging binding cooperativity. *J Med Chem.* 2022;65: 2940–2955.
- 77 Wertz IE, Murray JM. Structurally-defined deubiquitinase inhibitors provide opportunities to investigate disease mechanisms. *Drug Discov Today Technol.* 2019; 31:109–123.
- 78 Zhao Y, Du X, Duan Y, et al. High-throughput screening identifies established drugs as SARS-CoV-2 PLpro inhibitors. *Protein Cell.* 2021;12:877–888.
- 79 Ma C, Sacco MD, Xia Z, et al. Discovery of SARS-CoV-2 Papain-like Protease Inhibitors through a Combination of High-Throughput Screening and a FlipGFP-Based Reporter Assay. *ACS Cent Sci.* 2021;7:1245–1260.
- 80 Drag M, Mikolajczyk J, Bekes M, et al. Positional-scanning fluorogenic substrate libraries reveal unexpected specificity determinants of DUBs (deubiquitinating enzymes). *Biochem J.* 2008;415:367–375.
- 81 Rut W, Zmudzinski M, Snipas SJ, Bekes M, Huang TT, Drag M. Engineered unnatural ubiquitin for optimal detection of deubiquitinating enzymes. *Chem Sci.* 2020;11:6058–6069.
- 82 Swatek KN, Aumayr M, Pruneda JN, et al. Irreversible inactivation of ISG15 by a viral leader protease enables alternative infection detection strategies. *Proc Natl Acad Sci USA.* 2018;115:2371–2376.
- 83 Akutsu M, Ye Y, Virdee S, Chin JW, Komander D. Molecular basis for ubiquitin and ISG15 cross-reactivity in viral ovarian tumor domains. *Proc Natl Acad Sci USA.* 2011;108:2228–2233.
- 84 James TW, Frias-Staheli N, Bacik J-P, et al. Structural basis for the removal of ubiquitin and interferon-stimulated gene 15 by a viral ovarian tumor domain-containing protease. *Proc Natl Acad Sci USA.* 2011;108:2222–2227.
- 85 Huai W, Liu X, Wang C, et al. KAT8 selectively inhibits antiviral immunity by acetylating IRF3. *J Exp Med.* 2019;216:772–785.
- 86 Wang C, Wang Q, Xu X, et al. The methyltransferase NSD3 promotes antiviral innate immunity via direct lysine methylation of IRF3. *J Exp Med.* 2017;214: 3597–3610.
- 87 Munnur D, Teo Q, Eggermont D, et al. Altered ISGylation drives aberrant macrophage-dependent immune responses during SARS-CoV-2 infection. *Nat Immunol.* 2021;22:1416–1427.
- 88 Bade VN, Nickels J, Keusekotten K, Praefcke GJK. Covalent protein modification with ISG15 via a conserved cysteine in the hinge region. *PLoS One.* 2012;7:e38294.
- 89 Jeon SJ, Chung KC. Covalent conjugation of ubiquitin-like ISG15 to apoptosis-inducing factor exacerbates toxic stimuli-induced apoptotic cell death. *J Biol Chem.* 2022;298, 102464.
- 90 Durfee LA, Lyon N, Seo K, Huibregtse JM. The ISG15 conjugation system broadly targets newly synthesized proteins: implications for the antiviral function of ISG15. *Mol Cell.* 2010;38:722–732.
- 91 Zhao C, Hsiang T-Y, Kuo R-L, Krug RM. ISG15 conjugation system targets the viral NS1 protein in influenza A virus-infected cells. *Proc Natl Acad Sci USA.* 2010;107: 2253–2258.
- 92 Perng Y-C, Lenschow DJ. ISG15 in antiviral immunity and beyond. *Nat Rev Microbiol.* 2018;16:423–439.
- 93 Frias-Staheli N, Giannakopoulos NV, Kikkert M, et al. Ovarian tumor domain-containing viral proteases evade ubiquitin- and ISG15-dependent innate immune responses. *Cell Host Microbe.* 2007;2:404–416.
- 94 Isaacson MK, Ploegh HL. Ubiquitination, ubiquitin-like modifiers, and deubiquitination in viral infection. *Cell Host Microbe.* 2009;5:559–570.
- 95 Hung H-C, Wang H-C, Shih S-R, Teng IF, Tseng C-P, Hsu JTA. Synergistic inhibition of enterovirus 71 replication by interferon and rupintrivir. *J Infect Dis.* 2011;203: 1784–1790.
- 96 Xiang Z, Liu L, Lei X, Zhou Z, He B, Wang J. 3C protease of enterovirus D68 inhibits cellular defense mediated by interferon regulatory factor 7. *J Virol.* 2016;90: 1613–1621.
- 97 Lei X, Xiao X, Xue Q, Jin Q, He B, Wang J. Cleavage of interferon regulatory factor 7 by enterovirus 71 3C suppresses cellular responses. *J Virol.* 2013;87:1690–1698.
- 98 Zhao Y, Zhu Y, Liu X, et al. Structural basis for replicase polyprotein cleavage and substrate specificity of main protease from SARS-CoV-2. *Proc Natl Acad Sci USA.* 2022;119, e2117142119.
- 99 Pablos I, Machado Y, de Jesus HCR, et al. Mechanistic insights into COVID-19 by global analysis of the SARS-CoV-2 3CLpro substrate degradome. *Cell Rep.* 2021;37, 109892.
- 100 Koudelka T, Boger J, Henkel A, et al. N-Terminomics for the identification of in vitro substrates and cleavage site specificity of the SARS-CoV-2 main protease. *Proteomics.* 2021;21:2000246.
- 101 Yan W, Zheng Y, Zeng X, He B, Cheng W. Structural biology of SARS-CoV-2: open the door for novel therapies. *Sig Transduct Target Ther.* 2022;7:26.
- 102 Xu D, Shan B, Lee B-H, et al. Phosphorylation and activation of ubiquitin-specific protease-14 by Akt regulates the ubiquitin-proteasome system. *Elife.* 2015;4: e10510.
- 103 Kneller DW, Phillips G, O'Neill HM, et al. Structural plasticity of SARS-CoV-2 3CL Mpro active site cavity revealed by room temperature X-ray crystallography. *Nat Commun.* 2020;11:3202.

- 104 Lee J, Kenward C, Worrall LJ, et al. X-ray crystallographic characterization of the SARS-CoV-2 main protease polyprotein cleavage sites essential for viral processing and maturation. *Nat Commun.* 2022;13:5196.
- 105 Chan HTH, Moesser MA, Walters RK, et al. Discovery of SARS-CoV-2 Mpro peptide inhibitors from modelling substrate and ligand binding. *Chem Sci.* 2021;12:13686–13703.
- 106 Chan HTH, Oliveira ASF, Schofield CJ, Mulholland AJ, Duarte F. Dynamical nonequilibrium molecular dynamics simulations Identify allosteric sites and positions associated with drug resistance in the SARS-CoV-2 main protease. *JACS Au.* 2023;3:1767–1774.
- 107 Malla TR, Tumber A, John T, et al. Mass spectrometry reveals potential of  $\beta$ -lactams as SARS-CoV-2 Mpro inhibitors. *Chem Commun.* 2021;57:1430–1433.
- 108 Brewitz L, Dumjahn L, Zhao Y, et al. Alkyne derivatives of SARS-CoV-2 main protease inhibitors including nirmatrelvir inhibit by reacting covalently with the nucleophilic cysteine. *J Med Chem.* 2023;66:2663–2680.
- 109 Malla TR, Brewitz L, Muntean D-G, et al. Penicillin derivatives inhibit the SARS-CoV-2 main protease by reaction with its nucleophilic cysteine. *J Med Chem.* 2022;65:7682–7696.
- 110 Thun-Hohenstein STD, Suits TF, Malla TR, et al. Structure-activity studies reveal scope for optimisation of ebselen-type inhibition of SARS-CoV-2 main protease. *ChemMedChem.* 2022;17:e202100582.
- 111 Miura T, Malla TR, Owen CD, et al. In vitro selection of macrocyclic peptide inhibitors containing cyclic  $\gamma$ 2,4-amino acids targeting the SARS-CoV-2 main protease. *Nat Chem.* 2023;15:998–1005.
- 112 Mahjour B, Zhang R, Shen Y, et al. Rapid planning and analysis of high-throughput experiment arrays for reaction discovery. *Nat Commun.* 2023;14:3924.
- 113 de Munnik M, Lithgow J, Brewitz L, et al.  $\alpha,\beta,\alpha'\beta'$ -Diepoxyketones are mechanism-based inhibitors of nucleophilic cysteine enzymes. *Chem Commun.* 2023. <https://doi.org/10.1039/D3CC02932H>.
- 114 Schrödinger LLC., The PyMOL Molecular Graphics System, Version 2.3.0.
- 115 Dunbrack Jr RL, Cohen FE. Bayesian statistical analysis of protein side-chain rotamer preferences. *Protein Sci.* 1997;6:1661–1681.
- 116 Lindorff-Larsen K, Piana S, Palmo K, et al. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins.* 2010;78:1950–1958.
- 117 Abraham MJ, Murtola T, Schulz R, et al. GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX.* 2015;1–2:19–25.
- 118 Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. *J Chem Phys.* 1983;79:926–935.
- 119 Hoops SC, Anderson KW, Merz Jr KM. Force field design for metalloproteins. *J Am Chem Soc.* 1991;113:8262–8270.
- 120 Bayly CI, Cieplak P, Cornell WD, Kollman PA. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J Phys Chem.* 1993;97:10269–10280.
- 121 Cornell WD, Cieplak P, Bayly CI, Kollman PA. Application of RESP charges to calculate conformational energies, hydrogen bond energies, and free energies of solvation. *J Am Chem Soc.* 1993;115:9620–9631.
- 122 Cornell WD, Cieplak P, Bayly CI, et al. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J Am Chem Soc.* 1995;117:5179–5197.
- 123 Frisch MJ, Trucks GW, Schlegel HB, et al. *Gaussian 16 Rev. A.03*. CT: Wallingford; 2016.
- 124 Case DA, Ben-Shalom IY, Brozell SR, et al. *AMBER 2018*. San Francisco: University of California; 2018.
- 125 Bussi G, Donadio D, Parrinello M. Canonical sampling through velocity rescaling. *J Chem Phys.* 2007;126, 014101.
- 126 Nosé S, Klein ML. Constant pressure molecular dynamics for molecular systems. *Mol Phys.* 1983;50:1055–1076.
- 127 Parrinello M, Rahman A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J Appl Phys.* 1981;52:7182–7190.
- 128 Darden T, York D, Pedersen L. Particle mesh Ewald: an  $N$ -log( $N$ ) method for Ewald sums in large systems. *J Chem Phys.* 1993;98:10089–10092.
- 129 Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG. A smooth particle mesh Ewald method. *J Chem Phys.* 1995;103:8577–8593.
- 130 Daura X, Gademann K, Jaun B, Seebach D, van Gunsteren WF, Mark AE. Peptide folding: when simulation meets experiment. *Angew Chem Int Ed.* 1999;38:236–240.
- 131 Srinivasan J, Cheatham TE, Cieplak P, Kollman PA, Case DA. Continuum solvent studies of the stability of DNA, RNA, and phosphoramidate–DNA helices. *J Am Chem Soc.* 1998;120:9401–9409.
- 132 Kollman PA, Massova I, Reyes C, et al. Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Acc Chem Res.* 2000;33:889–897.
- 133 Gohlke H, Kiel C, Case DA. Insights into protein–protein binding by binding free energy calculation and free energy decomposition for the Ras-Raf and Ras-RalGDS complexes. *J Mol Biol.* 2003;330:891–913.
- 134 Metz A, Pflieger C, Kopitz H, Pfeiffer S, et al. Hot spots and transient pockets: predicting the determinants of small-molecule binding to a protein–protein interface. *J Chem Inf Model.* 2012;52:120–133.
- 135 Miller III BR, McGee Jr TD, Swails JM, Homeyer N, Gohlke H, Roitberg AE. MMPBSA.py: an efficient program for end-state free energy calculations. *J Chem Theory Comput.* 2012;8:3314–3321.
- 136 Onufriev A, Bashford D, Case DA. Exploring protein native states and large-scale conformational changes with a modified generalized born model. *Proteins.* 2004;55:383–394.