



Cite this: *Chem. Sci.*, 2020, **11**, 6058

All publication charges for this article have been paid for by the Royal Society of Chemistry

## Engineered unnatural ubiquitin for optimal detection of deubiquitinating enzymes†

Violetta Rut, <sup>a</sup> Mikolaj Zmudzinski, <sup>a</sup> Scott J. Snipas, <sup>b</sup> Miklos Bekes, <sup>†,c</sup> Tony T. Huang <sup>c</sup> and Marcin Drag <sup>ab</sup>

Deubiquitinating enzymes (DUBs) are responsible for removing ubiquitin (Ub) from its protein conjugates. DUBs have been implicated as attractive therapeutic targets in the treatment of viral diseases, neurodegenerative disorders and cancer. The lack of selective chemical tools for the exploration of these enzymes significantly impairs the determination of their roles in both normal and pathological states. Commercially available fluorogenic substrates are based on the C-terminal Ub motif or contain Ub coupled to a fluorophore (Z-LRGG-AMC, Ub-AMC); therefore, these substrates suffer from lack of selectivity. By using a hybrid combinatorial substrate library (HyCoSuL) and a defined P2 library containing a wide variety of nonproteinogenic amino acids, we established a full substrate specificity profile for two DUBs—MERS PLpro and human UCH-L3. Based on these results, we designed and synthesized Ub-based substrates and activity-based probes (ABPs) containing selected unnatural amino acids located in the C-terminal Ub motif. Biochemical analysis and cell lysate experiments confirmed the activity and selectivity of engineered Ub-based substrates and probes. Using this approach, we propose that for any protease that recognizes Ub and Ub-like substrates, a highly active and selective unnatural substrate or probe can be engineered.

Received 5th March 2020  
Accepted 27th May 2020

DOI: 10.1039/d0sc01347a  
[rsc.li/chemical-science](http://rsc.li/chemical-science)

## 1. Introduction

Ubiquitin (Ub) conjugation is one of the most important post-translational modifications and influences the activity, protein-protein interactions, localization and stability of many proteins.<sup>1</sup> Attachment of Ub to protein substrates is a covalent modification mediated by the E1, E2 and E3 enzymes. However, due to the activity of deubiquitinating enzymes (DUBs), this modification is reversible. DUBs hydrolyze the isopeptide bond between the C-terminal glycine of Ub and its substrate or between Ubs in poly-Ub chains with various topologies.<sup>2</sup> The human genome encodes approximately 100 DUBs divided into six subclasses based on sequence similarity and mechanism of action. The majority of DUBs (five of the six subclasses) are cysteine proteases, and only one subclass is represented by

a small group of metalloproteases. The largest and most structurally diverse subclass of DUBs are Ub-specific proteases (USPs). These enzymes can hydrolyze a peptide/isopeptide bond between the C-terminus of Ub and its substrate, as well as within the poly-Ub chain. Another subclass of DUBs that belongs to cysteine proteases are the Ub carboxy-terminal hydrolase (UCH) enzymes responsible for the hydrolysis of Ub precursors, leading to the formation of Ub monomers. The other subclasses of cysteine proteases are the ovarian tumor (OTU), Machado-Josephine domain (MJD), and motif-interacting with Ub-containing novel DUB family (MINDY) proteases and one metalloprotease subgroup, known as the JAMM/MPN+ DUBs. Members of the OTU family exhibit unique specificity for cleavage of different poly-Ub chain topologies.<sup>3</sup> Furthermore, enzymes displaying deubiquitinating activity are encoded in coronavirus genomes, such as Middle East respiratory syndrome papain-like protease (MERS PLpro)<sup>4</sup> or severe acute respiratory syndrome papain-like protease (SARS PLpro).<sup>5</sup> PLpro enzymes are involved in negative regulation of the innate immune response during viral infection.<sup>4</sup>

Since DUBs are highly specific toward Ub, the most commonly used DUB assay reagents are based on Ub with a C-terminal fluorogenic group (Ub-ACC,<sup>6</sup> Ub-AMC,<sup>7</sup> Ub-rhodamine<sup>8</sup>) or an electrophilic warhead (Ub-CHO, Ub-VS, Ub-vinyl methyl ester (VME)<sup>9–12</sup>). Application of these chemical tools in biological studies has led to a better understanding of the mechanisms of action of many DUBs. Ub with a C-terminal

<sup>a</sup>Department of Chemical Biology and Bioimaging, Wroclaw University of Science and Technology, Wyb. Wyspianskiego 27, 50-370 Wroclaw, Poland. E-mail: violetta.rut@pwr.edu.pl; marcin.drag@pwr.edu.pl

<sup>b</sup>Sanford Burnham Prebys Medical Discovery Institute, 10901 North Torrey Pines Road, La Jolla, CA 92037, USA

<sup>c</sup>Department of Biochemistry & Molecular Pharmacology, New York University School of Medicine, New York, NY 10016, USA

† Electronic supplementary information (ESI) available: DUBs substrate specificity, DUB labelling in cell lysates, pull-down experiment, structures of natural and unnatural amino acids and characterization of all compounds (PDF). See DOI: 10.1039/d0sc01347a

‡ Present address: Arvinas, Inc., 5 Science Park, New Haven, CT, 06511, USA.



aldehyde group has been frequently used in crystallographic studies of DUBs.<sup>13</sup> Development of Ub-based probes by introducing a detection tag on the N-terminus and a reactive group on the C-terminus of Ub enabled labeling of DUBs in cell lysates and their detection *via* western blot analysis. Ub-based probes are powerful tools for characterizing the activity of DUBs, identifying new DUBs, evaluating the selectivity of DUB inhibitors, and determining the functional roles of DUBs in normal and pathological states.<sup>12,14,15</sup> However, the use of mono-Ub probes precluded studies of DUB activity and linkage specificity. To overcome this limitation, di-Ub probes were developed.<sup>16,17</sup> Studies with di-Ub probes revealed differences between the S1 – S1' and S1 – S2 selectivity of enzymes from the OTU subclass,<sup>18</sup> as well as the Lys48 linkage specificity of SARS PLpro.<sup>19,20</sup> Despite the wide utility of broad-spectrum DUB substrates and probes in biochemical assays, selective chemical tools to study a single DUB (or a narrow subset of DUBs) are needed. Recently, the Ovaa group developed Ub-based probes selective toward USP7<sup>21</sup> and USP16<sup>22</sup> by combining structural analysis, modeling and mutational predictions. Selectivity can also be achieved by modifying the Ub C-terminal conserved LRGG motif.<sup>23</sup> Determination of DUB substrate preferences in the P4–P2 positions using the positional scanning substrate combinatorial library (PS-SCL) approach revealed that DUBs can recognize amino acids other than the canonical LRG in these positions. Furthermore, this study shows that a Ub mutant with C-terminal extension, in which glycine in the P2 position was replaced by valine, was cleaved by UCH-L3 20–50-fold less efficiently than its wild-type analog.<sup>23</sup> However, this mutation provided selectivity, as the mutant was not recognized by another tested DUB USP5 (IsoT). In the present study, to further improve Ub mutant selectivity, as well as DUB activity, we decided to utilize unnatural amino acids. A similar approach was successfully applied to design selective substrates and activity-based probes (ABPs) for cysteine,<sup>24,25</sup> serine<sup>26</sup> and threonine proteases<sup>27</sup> as short peptide sequences but never incorporated into whole substrate proteins, such as Ub.

In this work, we developed active and selective chemical tools based on Ub scaffolds for DUB investigation by incorporating unnatural amino acids into the conserved C-terminal Ub motif. We selected two DUBs, namely, UCH-L3 and MERS PLpro, to validate our chemical approach. These enzymes represent potential therapeutic targets in drug design.<sup>4,28,29</sup> First, we determined full substrate specificity profiles of the selected DUBs at the P4–P2 positions using defined and hybrid combinatorial libraries of tetrapeptide fluorogenic substrates. Substrate preference analysis allowed us to extract key amino acids that could serve as selectivity-providing motifs in Ub-based substrates and ABPs. The designed tetrapeptide fluorogenic substrates containing unnatural amino acids were selective toward UCH-L3 and MERS PLpro but still not efficiently hydrolyzed. To overcome this obstacle, we synthesized Ub derivatives containing a C-terminal epitope substituted with selected unnatural peptide sequences instead of a canonical LRGG motif. Total chemical synthesis of proteins provides flexibility in incorporation of different PTMs, warheads, non-canonical amino acids or other moieties in straightforward

manner.<sup>30</sup> Substrates and covalent probes were obtained by conjugation of the fluorescent label (ACC) or Michael-acceptor warhead (VME), respectively (Fig. 1). Kinetic analysis revealed that the designed substrates were more efficiently hydrolyzed than the reference “wild-type” substrate Ub-ACC. Moreover, the substrates exhibited high selectivity toward UCH-L3 and MERS PLpro. The selectivity of the Ub-based probe was evaluated using purified recombinant DUBs and cell lysates. The results of these studies revealed that engineered B-Ub-UCHL3-1-VME and B-Ub-UCHL3-2-VME probes detected UCH-L3 in cell lysates with high selectivity. With this *proof of concept*, we propose that this approach can be successfully applied in the design of Ub-based selective chemical tools for other proteases that recognize Ub and Ub-like substrates.

## 2. Experimental

### 2.1. Reagents

All chemical reagents were obtained from commercial suppliers and used without further purification. The reagents used for the solid-phase peptide synthesis (SPPS) were as follows: Rink Amide (RA) resin (particle size 100–200 mesh, loading 0.74 mmol g<sup>-1</sup>), 2-chlorotriptyl chloride resin (particle size 100–200 mesh, loading 0.97 mmol g<sup>-1</sup>), all Fmoc-amino acids, *O*-(benzotriazol-1-yl)-*N,N,N',N'*-tetramethyl-uronium-hexafluorophosphate (HBTU), 2-(1H-7-azabenzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HATU), piperidine, diisopropylcarbodiimide (DICl) and trifluoroacetic acid (TFA), purchased from Iris Biotech GmbH (Marktredwitz, Germany); Fmoc-Asp(OtBu)-(Dmb)Gly-OH and Fmoc-Ser(tBu)-Thr( $\psi$ Me,Me-pro)-OH (dipeptide building blocks) from AAPPTec (Louisville, KY, USA); 9-fluorenylmethyl carbazate (Fmoc-NH-NH<sub>2</sub>), tris(2-carboxyethyl)phosphine hydrochloride (TCEP) and 2,2'-azobis[2-(imidazolin-2-yl)propane] (VA-044) from Combi-Blocks (San Diego, CA, USA); 4-mercaptophenylacetic acid (MPAA), guanidine hydrochloride (Gn·HCl), sodium phosphate monobasic monohydrate (NaH<sub>2</sub>PO<sub>4</sub>·H<sub>2</sub>O) and 4.0 M HCl in dioxane; 2-methyl-2-propanethiol (tBuSH) from Sigma-Aldrich (Poznan, Poland); NaNO<sub>2</sub> from Merck Millipore (Warsaw, Poland); anhydrous *N*-hydroxybenzotriazole (HOBT) from Creosauls Louisville, KY, USA; 2,4,6-collidine (2,4,6-trimethylpyridine), HPLC-grade acetonitrile, triisopropylsilane (TIPS), tBu-N-allyl carbamate, toluene, methyl acrylate, dichlorophenylborane and 2<sup>nd</sup> generation Grubbs catalyst from Sigma-Aldrich (Poznan, Poland); and *N,N*-diisopropylethylamide (DIPEA) from VWR International (Gdansk, Poland). *N,N*-dimethylformamide (DMF), dichloromethane (DCM), methanol (MeOH), diethyl ether (Et<sub>2</sub>O), acetic acid (AcOH), and phosphorus pentoxide (P<sub>2</sub>O<sub>5</sub>), obtained from Avantor (Gliwice, Poland). Individual substrates, Ub-based substrates and probes were purified by HPLC on a Waters M600 solvent delivery module with a Waters M2489 detector system using a semipreparative Wide Pore C8 Discovery column and Jupiter 10  $\mu$ m C4 300  $\text{\AA}$  column (250  $\times$  10 mm). The solvent composition was as follows: phase A (water/0.1% TFA) and phase B (acetonitrile/0.1% TFA). The purity of each compound was confirmed with an analytical HPLC system using a Jupiter 10  $\mu$ m C4 300  $\text{\AA}$  column (250  $\times$  4.6 mm). The



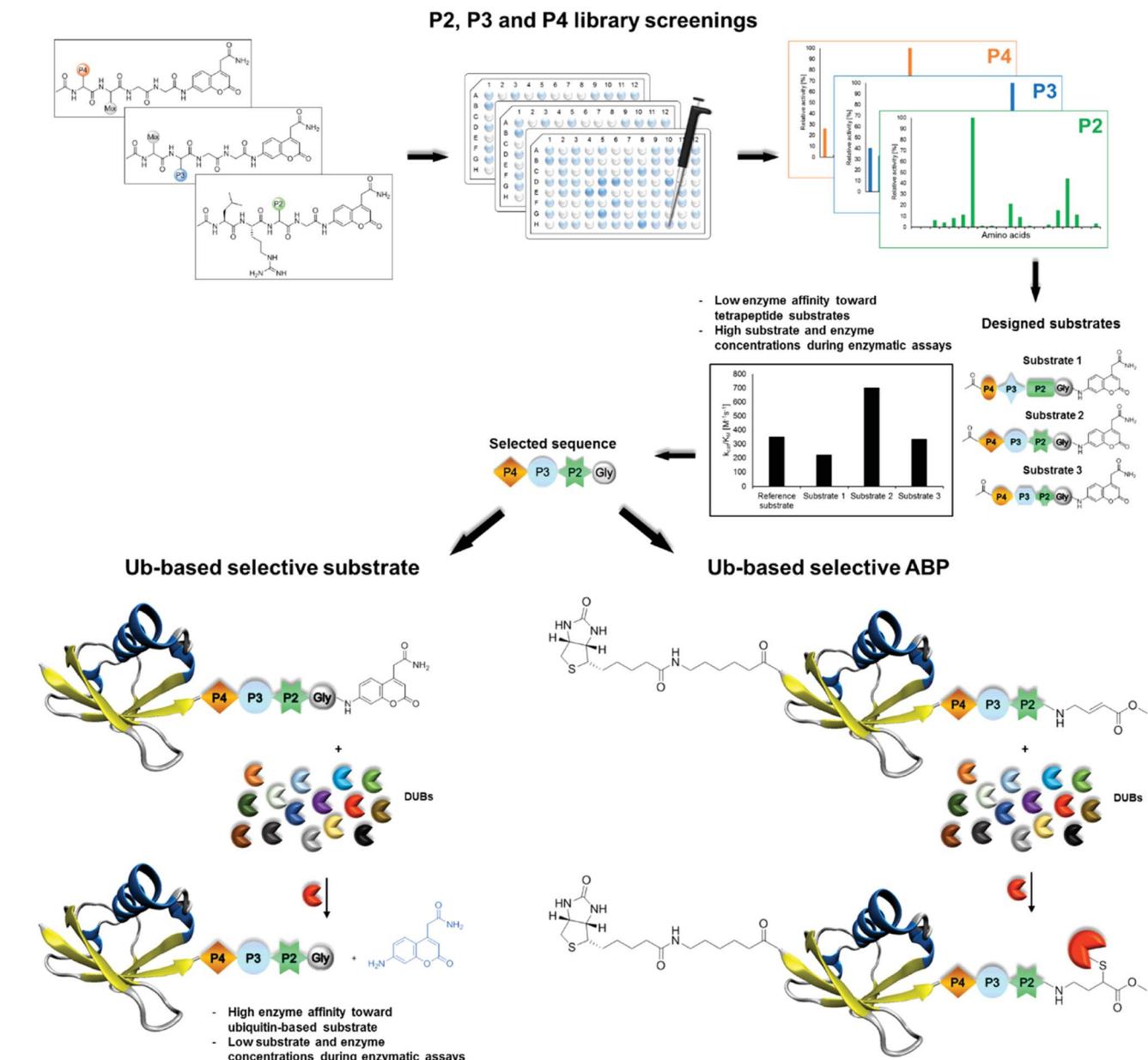


Fig. 1 Workflow applied in the design of DUB-selective substrates and ABPs (PDB ID: 1UBQ).

solvent composition was as follows: phase A (water/0.1% TFA) and phase B (acetonitrile/0.1% TFA); gradient, from 5% B to 95% B over a period of 15 or 20 min. The molecular weight of each substrate and ABP was confirmed by high-resolution mass spectrometry on a High-Resolution Mass Spectrometer Waters LCT premier XE with electrospray ionization (ESI) and a time-of-flight (TOF) module.

## 2.2. Combinatorial substrate library synthesis

A combinatorial tetrapeptide fluorogenic substrate library was synthesized on a solid support according to published protocols.<sup>31,32</sup> The library consisted of two tetrapeptide sublibraries. Each of the sublibraries was synthesized separately, and the general synthetic procedure is described for the P3 sublibrary as

an example. In the first step, Fmoc-ACC-OH (25 mmol, 2.5 eq.) was attached to the RA resin (13.5 g) using coupling reagents: HOEt (25 mmol, 2.5 eq.) and DICI (25 mmol, 2.5 eq.) in DMF. After 24 h, the Fmoc protecting group was removed with 20% piperidine in DMF. In the next step, Fmoc-Gly-OH (25 mmol, 2.5 eq.) was coupled using HATU (25 mmol, 2.5 eq.) and 2,4,6-collidine (25 mmol, 2.5 eq.) in DMF. Then, Fmoc-Gly-OH (25 mmol, 2.5 eq.) was attached to the H<sub>2</sub>N-Gly-ACC-resin using HOEt and DICI (25 mmol, 2.5 eq.) as coupling reagents. After glycine coupling, the Fmoc group was removed (20% piperidine in DMF), and the resin was washed with DCM and MeOH and dried over P<sub>2</sub>O<sub>5</sub>. Then, the dried resin was divided into 138 portions. To each portion of the H<sub>2</sub>N-Gly-Gly-ACC-resin, natural or unnatural amino acids were attached, and the Fmoc

protecting group was removed (20% piperidine in DMF). To provide equimolar substitution of each natural amino acid in the P4 position, an isokinetic mixture of Fmoc-protected amino acids was utilized. The last two steps of P3 sublibrary synthesis included N-terminal acetylation (solution of AcOH, HBTU and DIPEA) and cleavage of peptides from the resin using a TFA : H<sub>2</sub>O : TIPS (95% : 2.5% : 2.5%, % v/v/v) mixture. Finally, the sublibrary was precipitated in Et<sub>2</sub>O, dissolved in a mixture of acetonitrile and water, lyophilized and dissolved in biochemical grade DMSO at a concentration of 20 mM. The obtained sublibrary was used for kinetic studies without further purification. The P4 sublibrary was synthesized in the same manner.

### 2.3. Synthesis of the defined P2 library and individual substrates

The defined P2 library (Ac-Leu-Arg-P2-Gly-ACC, where P2 represents 128 natural and unnatural amino acids) and individual tetrapeptide fluorogenic substrates were synthesized on a solid support using the SPPS method as previously described.<sup>31,33</sup> Each substrate was purified by HPLC and analyzed using analytical HPLC and MS. The purity of each compound was  $\geq 95\%$ . The P2 library and individual substrates were dissolved at 10 mM in DMSO and stored at  $-80^{\circ}\text{C}$  until use.

### 2.4. Library and substrate screening

All screenings were performed using a spectrofluorometer (Molecular Devices Spectramax Gemini XPS) in 96-well plates (Corning). The release of ACC was monitored continuously for 40 min at the appropriate wavelength ( $\lambda_{\text{ex}} = 355\text{ nm}$ ,  $\lambda_{\text{em}} = 460\text{ nm}$ ). For the assay, 1  $\mu\text{L}$  of substrate in DMSO was used with 99  $\mu\text{L}$  of enzyme. The enzyme was incubated in assay buffer (MERS PLpro: 20 mM Tris, 150 mM NaCl, 5 mM DTT, pH 8.0;<sup>19</sup> UCH-L3: 50 mM HEPES, 0.5 mM EDTA, 5 mM DTT, pH 7.5<sup>23</sup>) for 30 min at  $37^{\circ}\text{C}$  before addition to the substrates on a plate. The final substrate concentration in each well during the assays was 200  $\mu\text{M}$  for combinatorial P3 and P4 sublibraries, 100  $\mu\text{M}$  for the Ac-Leu-Arg-P2-Gly-ACC library and 10  $\mu\text{M}$  for individual substrates. The enzyme concentration was 2–3  $\mu\text{M}$  for MERS PLpro and 6–20  $\mu\text{M}$  for UCH-L3. Kinetic assays were repeated at least three times, and the results are presented as the mean values with standard deviations. The linear part of each progress curve was used to determine the substrate hydrolysis rate. Substrate specificity profiles were established by setting the highest value of relative fluorescence unit per second (RFU  $\text{s}^{-1}$ ) from each library position as 100% and adjusting other values accordingly.

### 2.5. Synthesis of Ub derivatives

Currently, different methods can be applied to obtain Ub and poly-Ub derivatives.<sup>6,34–36</sup> We modified existing strategies to develop a new, versatile approach that enables the synthesis of Ub-based substrates and ABPs containing unnatural amino acids. We employed this method to synthesize four different Ub-ACC substrates and four corresponding N-terminally biotinylated ABPs with a C-terminal VME electrophile. Each

substrate and probe was synthesized by applying a peptide hydrazide ligation strategy. Peptide hydrazide segments were synthesized by SPPS, purified by HPLC and assembled together in a stepwise manner. Then, we performed radical desulfurization and conjugated H<sub>2</sub>N-Gly-ACC to obtain Ub-ACC substrates or H<sub>2</sub>N-Gly-VME to obtain biotin-6-Ahx-Ub-VME ABPs. Herein, we present a general protocol for Ub derivative synthesis based on the example of Ub-ACC. We started the synthesis with Fmoc-hydrazide resin preparation based on a modified published method.<sup>37</sup> 2-CTC resin (0.97 mmol g<sup>-1</sup>) was swelled for 20 min in fresh DCM and washed with 3  $\times$  DCM. Fmoc-NH-NH<sub>2</sub> (4 eq.) was dissolved in DMF : DCM (8/1, v/v), and DIPEA (10 eq.) was added. The mixture was added to the resin. The reaction was carried out at room temperature. After 1 h, DIPEA (5 eq.) was added. Fmoc-NH-NH<sub>2</sub> coupling was carried out for 4 h. The resin was washed with 3  $\times$  DMF. End-capping of residual active 2-CTC resin was performed for 30 min by adding a capping mixture (DMF/MeOH/DIPEA, 3.5/0.5/0.2, v/v/v). After capping, Fmoc-NH-NH-2-CTC resin was washed with 3  $\times$  DMF. Ub derivatives were assembled from 3 peptide hydrazide segments: Ub[1–27]-NH-NH<sub>2</sub>, Ub[28–45]-(A<sup>28</sup>C)-NH-NH<sub>2</sub> and Ub[46–75]-(A<sup>46</sup>C)-NH-NH<sub>2</sub>. During the synthesis of ABPs, the Biot-6-Ahx-Ub[1–27]-NH-NH<sub>2</sub> segment was used instead of Ub[1–27]-NH-NH<sub>2</sub>. Naturally occurring Ala residues on the N-termini of Ub[28–45]-(A<sup>28</sup>C)-NH-NH<sub>2</sub> and Ub[46–75]-(A<sup>46</sup>C)-NH-NH<sub>2</sub> were exchanged to Cys residues to enable subsequent native chemical ligation (NCL) reactions. Segments were synthesized manually in a straightforward manner using Fmoc-protected amino acids. The synthesis of Ub[46–75]-(A<sup>46</sup>C)-NH-NH<sub>2</sub> required the use of dipeptide building blocks at Asp<sup>52</sup>Gly<sup>53</sup> (Fmoc-Asp(OtBu)-(Dmb)Gly-OH) and Ser<sup>65</sup>Thr<sup>66</sup> (Fmoc-Ser(tBu)-Thr( $\Psi^{\text{Me},\text{Me}}$ pro)-OH), as reported elsewhere.<sup>34</sup> Fmoc-based SPPS of peptide hydrazides was performed in glass peptide synthesis vessels. Fmoc-NH-NH-2-CTC resin was swelled for 20 min in DCM, deprotected with 20% piperidine in DMF (two cycles: 5 and 25 min) and washed with 7  $\times$  DMF. The first amino acid (2.5 eq.) was dissolved in DMF with coupling reagents (2.5 eq. HATU and 2.5 eq. 2,4,6-collidine). The amino acid was preactivated for 30 s and added to the resin. The mixture was gently agitated for 24 h. After coupling, the resin was washed with 3  $\times$  DMF, and the Fmoc groups were removed as described previously. A solution of the next amino acid (2.5 eq.) with HATU (2.5 eq.) and 2,4,6-collidine (2.5 eq.) in DMF was added to the resin. After complete coupling of the second amino acid, Fmoc deprotection was carried out, and the following amino acids or dipeptide building blocks were coupled with the same coupling-deprotection cycles. The progress of the coupling and Fmoc deprotection of primary amine groups was monitored with the ninhydrin test. If needed, double coupling was performed. Biotin in the Biot-6-Ahx-Ub[1–27]-NH-NH<sub>2</sub> segment was coupled in DMF/DMSO (1 : 1, v/v). After the last coupling step, the resin was washed with 3  $\times$  DMF, 3  $\times$  DCM, and 3  $\times$  MeOH and dried over P<sub>2</sub>O<sub>5</sub> overnight. A cleavage cocktail (8 mL, TFA/H<sub>2</sub>O/TIPS, 95/2.5/2.5, % v/v/v) was added to the dry resin. After 2 h, the cleavage mixture was filtered from the resin and collected.



Peptides were precipitated with cold  $\text{Et}_2\text{O}$  and dried. Crude peptide hydrazides were purified by preparative HPLC and lyophilized. NCL reactions were performed based on a published protocol.<sup>38</sup> Ub[1-27]-NH-NH<sub>2</sub> (7.13  $\mu\text{mol}$ ) was dissolved in ligation buffer (1.5 mL, 6 M  $\text{Gn}\cdot\text{HCl}$ , 0.2 M  $\text{NaH}_2\text{PO}_4$ , pH 3.0–3.1) and cooled in an ice-salt bath to  $-15\text{ }^\circ\text{C}$  for 20 min with stirring. Acyl azide generation was performed at  $-15\text{ }^\circ\text{C}$  for 20 min by the addition of  $\text{NaNO}_2$  (10 eq.). Then, 1 mL of a solution of Ub[28-45]-(A<sup>28</sup>C)-NH-NH<sub>2</sub> (1 eq.) and MPAA (100 eq.) in ligation buffer (pH adjusted to 6.5) was added to Ub[1-27]-N<sub>3</sub>. The ice-salt bath was removed, and when the reaction mixture reached room temperature, the pH was slowly adjusted to 6.8. The progress of the reaction was monitored by LC-MS. After completion of the reaction, reducing agent (TCEP, 200  $\mu\text{L}$ , 0.1 M in ligation buffer, pH 6.5) was added. The ligation product (Ub[1-45]-(A<sup>28</sup>C)-NH-NH<sub>2</sub>) was purified by semipreparative HPLC and lyophilized. Ligation of Ub[1-45]-(A<sup>28</sup>C)-NH-NH<sub>2</sub> with Ub[46-75]-(A<sup>46</sup>C)-NH-NH<sub>2</sub> was performed in the manner described above. Next, the product of the second ligation (Ub[1-75]-(A<sup>28</sup>C,A<sup>46</sup>C)-NH-NH<sub>2</sub>) was desulfurized based on a previously published protocol.<sup>6</sup> Ub[1-75]-(A<sup>28</sup>C,A<sup>46</sup>C)-NH-NH<sub>2</sub> (14.1 mg, 1.39  $\mu\text{mol}$ ) was dissolved in desulfurization buffer (1.8 mL, 6 M  $\text{Gn}\cdot\text{HCl}$ , 100 mM  $\text{Na}_2\text{HPO}_4$ , 375 mM TCEP, 37.5 mM VA-044, 150 mM *t*BuSH, pH 6.9). Desulfurization was performed for 4 h at  $37\text{ }^\circ\text{C}$ . The progress of the reaction was monitored by LC-MS. Ub[1-75]-NH-NH<sub>2</sub> was isolated by semipreparative HPLC and lyophilized. The last step of synthesis was  $\text{H}_2\text{N}\text{-Gly-ACC}$  conjugation. TFA· $\text{H}_2\text{N}\text{-Gly-ACC}$  was synthesized on RA resin (0.74 mmol g<sup>-1</sup>) and used without further purification. Ub[1-75]-NH-NH<sub>2</sub> (6.7 mg) was dissolved in DMF : DMSO (0.75 mL, 2 : 1, v/v). The mixture was cooled in an ice-salt bath to  $-20\text{ }^\circ\text{C}$  for 20 min with stirring.  $\text{NaNO}_2$  (10 eq.) in DMF : DMSO (50  $\mu\text{L}$ , 2 : 1, v/v) was added. The pH of the reaction mixture was adjusted to 3.0–4.0 with 4 M HCl in dioxane solution. Acyl azide generation was carried out for 30 min in an ice-salt bath, and the progress of the reaction was monitored by LC-MS. TFA· $\text{H}_2\text{N}\text{-Gly-ACC}$  (20 eq.) in DMF : DMSO (100  $\mu\text{L}$ , 2 : 1, v/v) and DIPEA (20 eq.) were added to the acyl azide, and the pH was adjusted to 8.0–9.0 with DIPEA. The reaction was performed at  $-20\text{ }^\circ\text{C}$  for 1 h. The progress of the reaction was monitored by LC-MS. The product was purified by analytical HPLC (purity > 95%), lyophilized, and characterized by HRMS. Synthesis of ABPs was performed in the same manner. The warhead synthesis protocol was adopted from a published method.<sup>39</sup> *t*Bu-N-allyl carbamate (500 mg, 3.2  $\mu\text{mol}$ ) was dissolved in 10 mL of anhydrous toluene. Methyl acrylate (580  $\mu\text{L}$ , 6.4  $\mu\text{mol}$ ), dichlorophenylborane (42  $\mu\text{L}$ , 0.32  $\mu\text{mol}$ ) and 2<sup>nd</sup> generation Grubbs catalyst (50 mg) were added. The reaction was carried out under reflux at  $40\text{ }^\circ\text{C}$  with stirring overnight. After 12 h the solvent was removed and the mixture was purified by column chromatography on silica gel (Hex/EtOAc 5 : 1). The crude product was obtained in the form of a yellowish oil. *t*Bu group deprotection was performed by adding TFA/DCM/TIPS (4.2 mL, 3/1/0.2, v/v/v) cleavage mixture for 45 min with stirring. TFA· $\text{H}_2\text{N}\text{-Gly-VME}$  was then crystallized in cold  $\text{Et}_2\text{O}$  and stored at  $-20\text{ }^\circ\text{C}$ .

## 2.6. Determination of kinetic parameters for tetrapeptide and Ub-based substrates

Kinetic studies were carried out in 96-well plates. Wells contained 20  $\mu\text{L}$  of ACC-labeled substrate at eight different concentrations (0.1–75  $\mu\text{M}$ ) and 80  $\mu\text{L}$  of enzyme in the same assay buffer as that described above (10 nM to 2.5  $\mu\text{M}$  MERS PLpro; 1 nM to 6  $\mu\text{M}$  UCH-L3). Substrate hydrolysis was measured for 30 min using the following wavelengths:  $\lambda_{\text{ex}} = 355\text{ nm}$ ,  $\lambda_{\text{em}} = 460\text{ nm}$ . Each experiment was repeated at least three times. Kinetic parameters were calculated using GraphPad Prism software with the Michaelis–Menten equation. Due to the precipitation of tetrapeptide substrates at high concentrations, only the specificity constant ( $k_{\text{cat}}/K_{\text{M}}$ ) was determined. When  $[S_0] \ll K_{\text{M}}$ , the plot of  $v_i$  (the initial velocities) versus  $[S_0]$  yields a straight line with slope representing  $V_{\text{max}}/K_{\text{M}}$ ,  $k_{\text{cat}}/K_{\text{M}} = \text{slope}/E$  ( $E$  – total enzyme concentration).

## 2.7. Determination of DUB inhibition by Ub-based probes

To assess DUB inhibition by biotinylated Ub-based probes, recombinant enzymes (50 nM) in assay buffer were incubated with eight different probe concentrations (0–3  $\mu\text{M}$ ) for 30 min at  $37\text{ }^\circ\text{C}$ . Then, Ub-ACC (30  $\mu\text{M}$ ) was added to estimate residual DUB activity. The release of ACC was monitored for 30 min at the following wavelengths:  $\lambda_{\text{ex}} = 355\text{ nm}$  and  $\lambda_{\text{em}} = 460\text{ nm}$ .

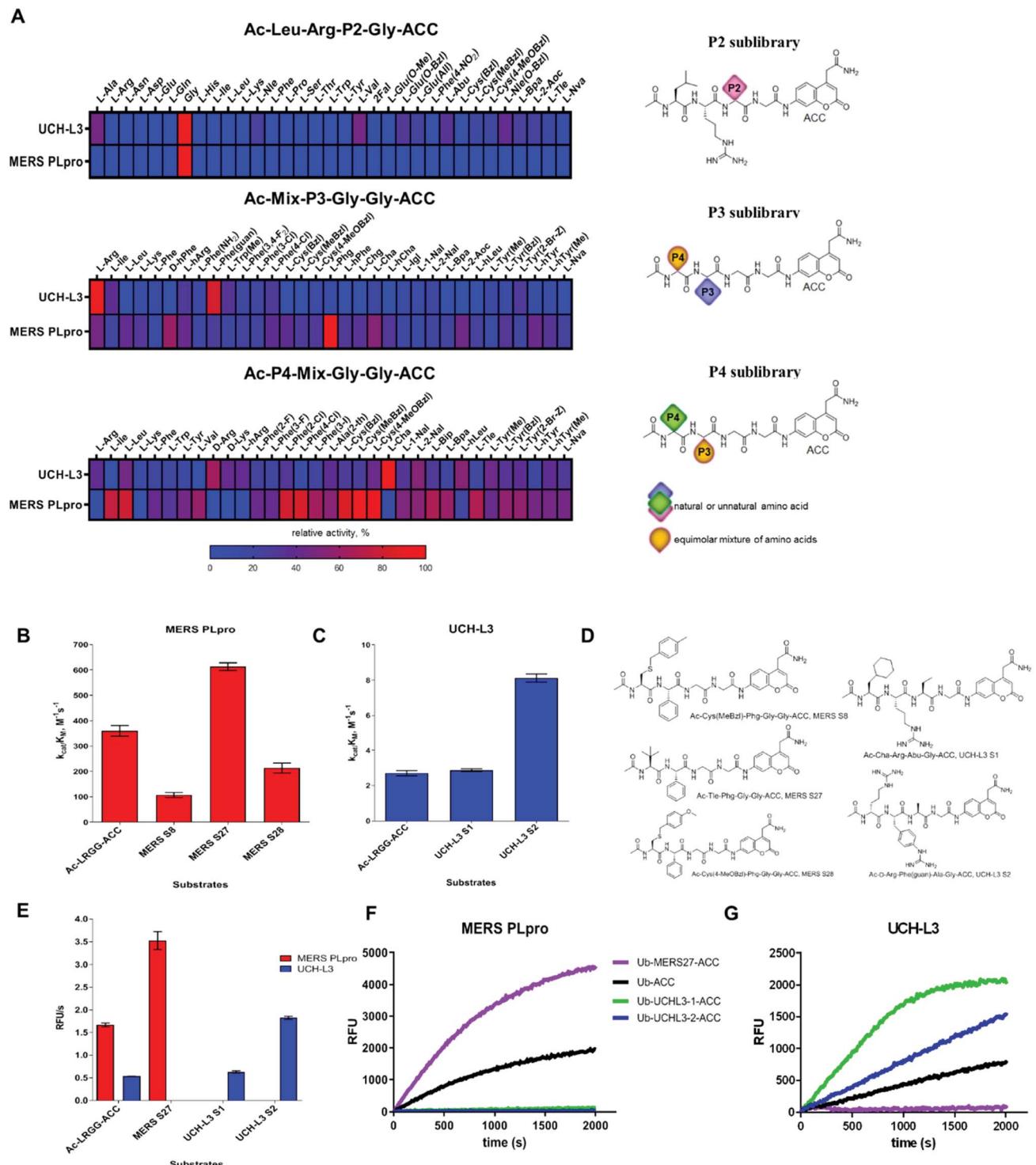
## 2.8. Determination of inhibition kinetics ( $k_{\text{obs}}/I$ ) for Ub-based probes

MERS PLpro (10 nM, enzyme was incubated in assay buffer 20 mM Tris, 150 mM NaCl, 5 mM DTT, pH 8.0 for 30 min at  $37\text{ }^\circ\text{C}$  before addition to a plate) was added to seven different wells with probe concentration ranging from 1.8  $\mu\text{M}$  to 7  $\mu\text{M}$  and 25  $\mu\text{M}$  of substrate (Ub-ACC). The measurement was conducted for 30 minutes and repeated at least three times,  $k_{\text{obs}}/I$  was calculated as previously described.<sup>40</sup>

## 2.9. DUB labeling in cell lysates

A-431, HeLa and HEK-293T cells were cultured in DMEM supplemented with 10% fetal bovine serum, 2 mM L-glutamine, and antibiotics (100 U mL<sup>-1</sup> penicillin, 100  $\mu\text{g mL}^{-1}$  streptomycin) in a humidified 5%  $\text{CO}_2$  atmosphere at  $37\text{ }^\circ\text{C}$ . Approximately 1 200 000 cells were harvested and washed three times with PBS. The cell pellet was lysed in buffer containing 20 mM Tris, 150 mM NaCl, and 5 mM DTT, pH 8.0, using a sonicator. The cell lysate was centrifuged for 10 min, and the supernatant was collected. Twenty microliters of lysate was incubated with 80  $\mu\text{L}$  of Ub-based probes at different concentrations for 30 min at  $37\text{ }^\circ\text{C}$ . Then, 50  $\mu\text{L}$  of 3× SDS/DTT was added, and the samples were boiled for 5 min at 95  $^\circ\text{C}$  and resolved on 4–12% Bis-Tris Plus 12-well gels at 30  $\mu\text{L}$  sample per well. Electrophoresis was performed at 200 V for 29 min or 165 V for 45 min. Next, the proteins were transferred to a nitrocellulose membrane (0.2  $\mu\text{m}$ , Bio-Rad) for 60 min at 10 V. The membrane was blocked with 2% BSA in Tris-buffered saline with 0.1% (v/v) Tween 20 (TBS-T) for 60 min at RT. Biotinylated Ub-based probes were detected with





**Fig. 2** Substrate specificity of UCH-L3 and MERS PLpro. (A) Substrate specificity profiles of MERS PLpro and UCH-L3 presented as heat maps (left panel) and structures of libraries (right panel). Kinetic parameters for tetrapeptide fluorogenic substrates with MERS PLpro (B) and UCH-L3 (C). (D) Tetrapeptide substrate structures: MERS S8, Ac-Cys(MeBzl)-Phg-Gly-Gly-ACC; MERS S27, Ac-Tle-Phg-Gly-Gly-ACC; MERS S28, Ac-Cys(4-MeOBz)-Phg-Gly-Gly-ACC; UCH-L3 S1, Ac-Cha-Arg-Abu-Gly-ACC; UCH-L3 S2, Ac-d-Arg-Phe(guan)-Ala-Gly-ACC. (E) Rate of tetrapeptide substrate hydrolysis by DUBs ( $[S] = 10 \mu\text{M}$ ; MERS PLpro concentration,  $2.5 \mu\text{M}$ ; UCH-L3 concentration,  $6 \mu\text{M}$ ); (F and G) Ub-based substrate selectivity ( $[S] = 1 \mu\text{M}$ ; MERS PLpro concentration,  $5 \text{nM}$ ; UCH-L3 concentration,  $1 \text{nM}$ ).

a fluorescent streptavidin Alexa Fluor 647 conjugate ( $1 : 10\,000$ ) in TBS-T with  $1\%$  BSA, and UCH-L3 was detected with a mouse anti-human monoclonal IgG<sub>1</sub> antibody ( $1 : 1000$ )

and fluorescent goat anti-mouse ( $1 : 10\,000$ ) using an Azure Biosystems Sapphire Biomolecular Imager and Azure Spot Analysis Software.

## 2.10. Pull-down experiment

Cell lysates were prepared in the same manner as described above. Cell lysates were incubated with or without Ub-based probes (100 nM) for 30 min at 37 °C. Before the addition of the beads (Thermo Fischer Scientific, 30353, washed multiple times with DPBS), 100 μM E-64 was added to the samples to inhibit protease activity. Streptavidin beads were added to the lysates and the samples were incubated at 4 °C for 2.5 h. Next, beads were collected by centrifugation at 1000 rpm, washed 5 times with DPBS, boiled at 95 °C for 15 min in loading buffer and analyzed by gel electrophoresis and western blotting.

## 3. Results

### 3.1. DUB substrate specificity profile

To date, the substrate specificity profile of selected DUBs has been determined using only natural amino acids.<sup>19,23</sup> It has been found that these enzymes can also recognize other amino acids at the P4-P2 positions, not only the C-terminal LRGG motif. These results suggest that dissection of the binding pocket architecture of DUBs can lead to the development of new chemical tools for DUB investigation with high specificity. To precisely examine the binding pocket preferences of selected enzymes, we used substrate libraries containing natural and a large number of unnatural amino acids with diverse chemical structures. Since all DUBs recognize leucine and arginine at the P4 and P3 positions respectively, we synthesized a defined library with a general structure of Ac-Leu-Arg-P2-Gly-ACC (where P2 is a natural or unnatural amino acid) to determine the substrate specificity profile of DUBs at the P2 position. To examine substrate preferences at the P4 and P3 positions, we designed and synthesized a hybrid combinatorial substrate library (HyCoSuL). This library consists of two sublibraries: Ac-P4-Mix-Gly-Gly-ACC and Ac-Mix-P3-Gly-Gly-ACC (where P4 and

P3 represent 19 natural and 119 unnatural amino acids, and Mix represents an equimolar mixture of natural amino acids). We incorporated glycine at the P1 and P2 positions because this amino acid is preferred by all DUBs in the S1 and S2 pockets. These two libraries can be used to determine the substrate specificity profile at the P4-P2 positions of all enzymes that display deubiquitinating activity.

**P2 position.** DUBs exhibited very narrow substrate specificity at the P2 position (Fig. 2A, see full substrate specificity profile in the ESI, Fig. S1†). MERS PLpro recognized only glycine at this position. The S2 pocket of UCH-L3 could accommodate not only glycine as the best hit but also, to a lesser extent, some aliphatic amino acids, such as Ala (47%), Val (46.5%), Abu (33.5%), Nle (25%), 2-Aoc (24%), and Tle (15%), and large hydrophobic amino acids, such as Nle(OBzl) (38%), Glu(OBzl) (31%), and Bpa (25%).

**P3 and P4 positions.** UCH-L3 displayed a narrow substrate preference at the P3 position (Fig. 2A and S1†). Positively charged amino acids were preferred by the S3 pocket (Arg (100%), Phe(guan) (81%), and hArg (30%)). Hydrophobic residues were also recognized but with lower affinity (<30%, except Trp(Me) (41%) and Ile (36%)). In contrast, MERS PLpro preferred hydrophobic amino acids at the P3 position over basic residues (Phg (100%), D-hPhe (60%), Cha (54%), Arg (47%), 2-Aoc (44%), hArg (42%), hTyr (42%), and Leu (38.5%); amino acid structures presented in Table S2 ESI†). The S4 pocket of UCH-L3 could accommodate hydrophobic amino acids (Cha (100%), Met (66.5%), hLeu (54%), 2-Nal (52.5%), Leu (46%), Cys(4-MeOBzl) (44%)) as well as positively charged L- and D-amino acids (D-Arg (64%), D-Lys (40.5%), hArg (39%), Arg (38%)). MERS PLpro favored aliphatic and bulky hydrophobic residues at the P4 position. The best recognized amino acids were Cys(MeBzl) (100%), Cys(4-MeOBzl) (93%), Cys(Bzl) (84%), Phe(4-Cl) (83%), Phe(2-Cl) (80%), Leu (79%), and Ile (75%). Basic and

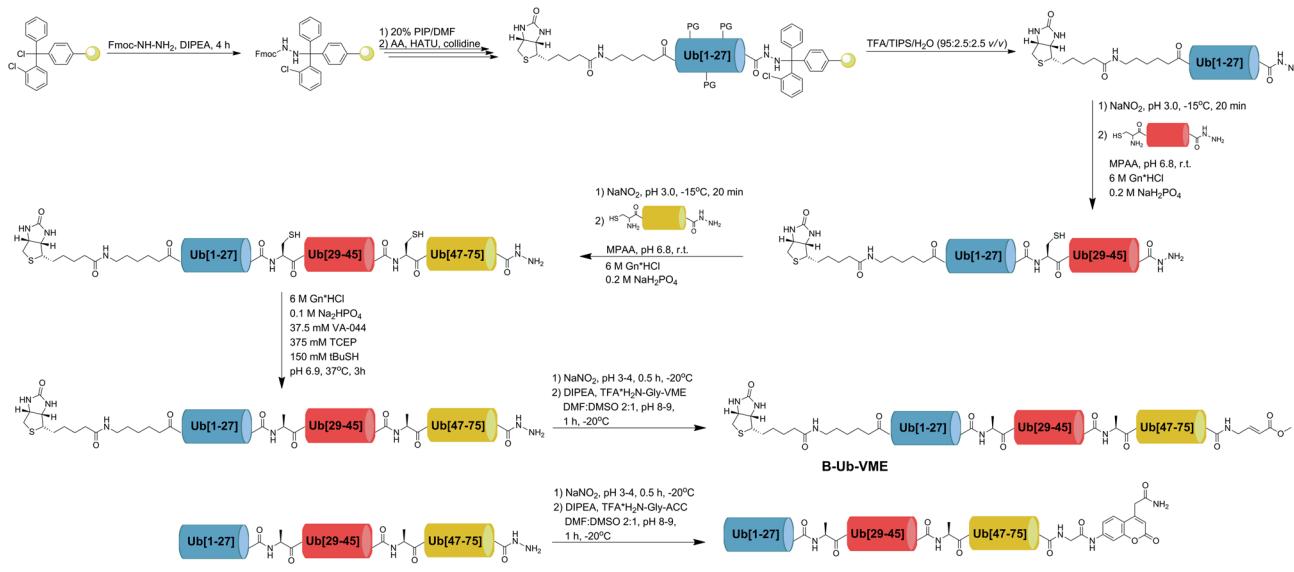


Fig. 3 Scheme of Ub derivative synthesis.



Table 1 Kinetic parameters of Ub-based fluorogenic substrates with UCH-L3 and MERS PLpro<sup>a</sup>

Ub-based substrate	Enzyme	$K_M$ , $\mu\text{M}$	$k_{\text{cat}}$ , $\text{s}^{-1}$	$k_{\text{cat}}/K_M$ , $\text{M}^{-1} \text{s}^{-1}$
Ub-ACC	UCH-L3	$0.048 \pm 0.002$	$0.107 \pm 0.004$	$(2.2 \pm 0.093) \times 10^6$
	MERS PLpro	$3.55 \pm 0.09$	$3.679 \pm 0.179$	$(1.0 \pm 0.034) \times 10^6$
Ub-MERS27-ACC	UCH-L3	$2.11 \pm 0.07$	$0.023 \pm 0.0005$	$(1.1 \pm 0.060) \times 10^4$
	MERS PLpro	$7.62 \pm 0.51$	$12.386 \pm 0.341$	$(1.6 \pm 0.076) \times 10^6$
Ub-UCHL3-1-ACC	UCH-L3	$0.099 \pm 0.011$	$0.387 \pm 0.017$	$(3.9 \pm 0.22) \times 10^6$
	MERS PLpro	ND	ND	ND
Ub-UCHL3-2-ACC	UCH-L3	$0.083 \pm 0.005$	$0.224 \pm 0.018$	$(2.7 \pm 0.25) \times 10^6$
	MERS PLpro	—	—	—

<sup>a</sup> “ND” not determined due to substrate precipitation at high concentration; “—” not recognized.

acidic residues were almost not recognized by MERS PLpro at this position.

### 3.2. Kinetic analysis of tetrapeptide fluorogenic substrates

To select the optimal peptide sequence for MERS PLpro and UCH-L3, we analyzed substrate specificity preferences at the P4–P2 positions for each enzyme and selected amino acids that were both well recognized and selective toward the tested DUBs. For MERS PLpro, we chose Phg at the P3 position as the best recognized and most selective amino acid. At the P4 position, we selected Cys(MeBzl), Cys(4-MeOBzl) and Tle as the most promising candidates. The S2 pocket of UCH-L3 could accommodate some aliphatic amino acids; thus, we decided to incorporate Ala and Abu residues in substrate sequences. Although these amino acids were less well recognized than glycine, they can have a significant effect on substrate selectivity. At the P3 and P4 positions, we selected certain amino acids as the best hits (P3: Arg, Phe(guan); P4: d-Arg, Cha). After selection of the amino acids, we synthesized ACC-labeled substrates and determined the catalytic efficiency of the enzyme (Fig. 2B–D). Most likely due to steric hindrance, MERS S8 and MERS S28 were not better hydrolyzed by MERS PLpro than the reference substrate (Ac-Leu-Arg-Gly-Gly-ACC). Ac-d-Arg-Phe(guan)-Ala-Gly-ACC (UCH-L3 S2) was 3 times better recognized by UCH-L3 than Ac-Leu-Arg-Gly-Gly-ACC. In the next step, we investigated substrate selectivity (Fig. 2E). Kinetic analysis revealed that Ac-Tle-Phg-Gly-Gly-ACC (MERS S27) was not recognized by UCH-L3, and Ac-d-Arg-Phe(guan)-Ala-Gly-ACC

and Ac-Cha-Arg-Abu-Gly-ACC (UCH-L3 S1) were selective toward this protease. We selected these three peptide sequences for further analysis.

### 3.3. Synthesis and kinetic analysis of fluorescent Ub derivative substrates

Since tetrapeptide fluorogenic substrates are not efficiently hydrolyzed by DUBs, even after incorporation of unnatural amino acids, the kinetic rates of the substrates were only 2–3 times higher compared to that of the reference Ac-LRGG-ACC substrate. Therefore, we decided to synthesize fluorescent Ub derivatives containing unnatural amino acids on the C-terminal tetrapeptidic epitope. To date, several chemical synthesis methods for Ub derivatives have been reported.<sup>6,34–36</sup> We applied these methods with several modifications for effective synthesis of Ub-based substrates and ABPs comprising unnatural building blocks. In the first step, we divided Ub into three peptide segments and synthesized them separately using Fmoc-based SPPS of peptide hydrazides on 2-chlorotriptyl chloride resin. This approach enables (1) efficient synthesis of each peptide segment with good isolated yields and purity; (2) incorporation of unnatural amino acids on the C-terminal Ub motif; and (3) modification of the N-terminus by introducing tags and linkers. In the next step, three peptide segments were assembled in a stepwise manner through NCL. Then, cysteine residues, which were introduced in the Ub sequence to enable NCL reactions, were converted to alanine using a free radical desulfurization reaction. In the last step, a fluorescent tag (ACC)

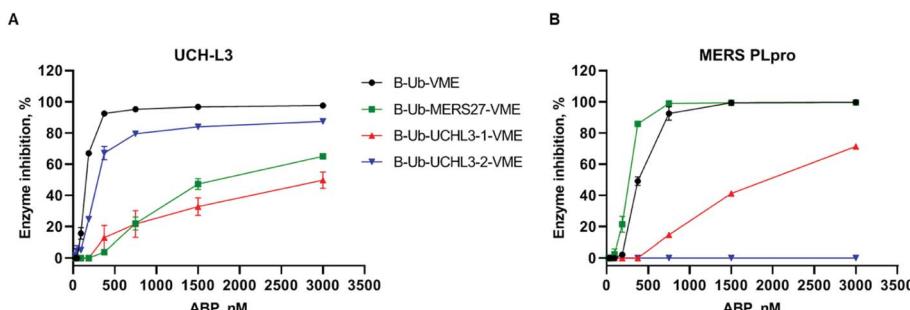


Fig. 4 Kinetic analysis of DUB inhibition by Ub probes. Recombinant enzymes (A) UCH-L3, (B) MERS PLpro were incubated with various probe concentrations for 30 min at 37 °C before DUB activity measurement. The experiment was performed in triplicate, and the average results are presented (SD < 10%).



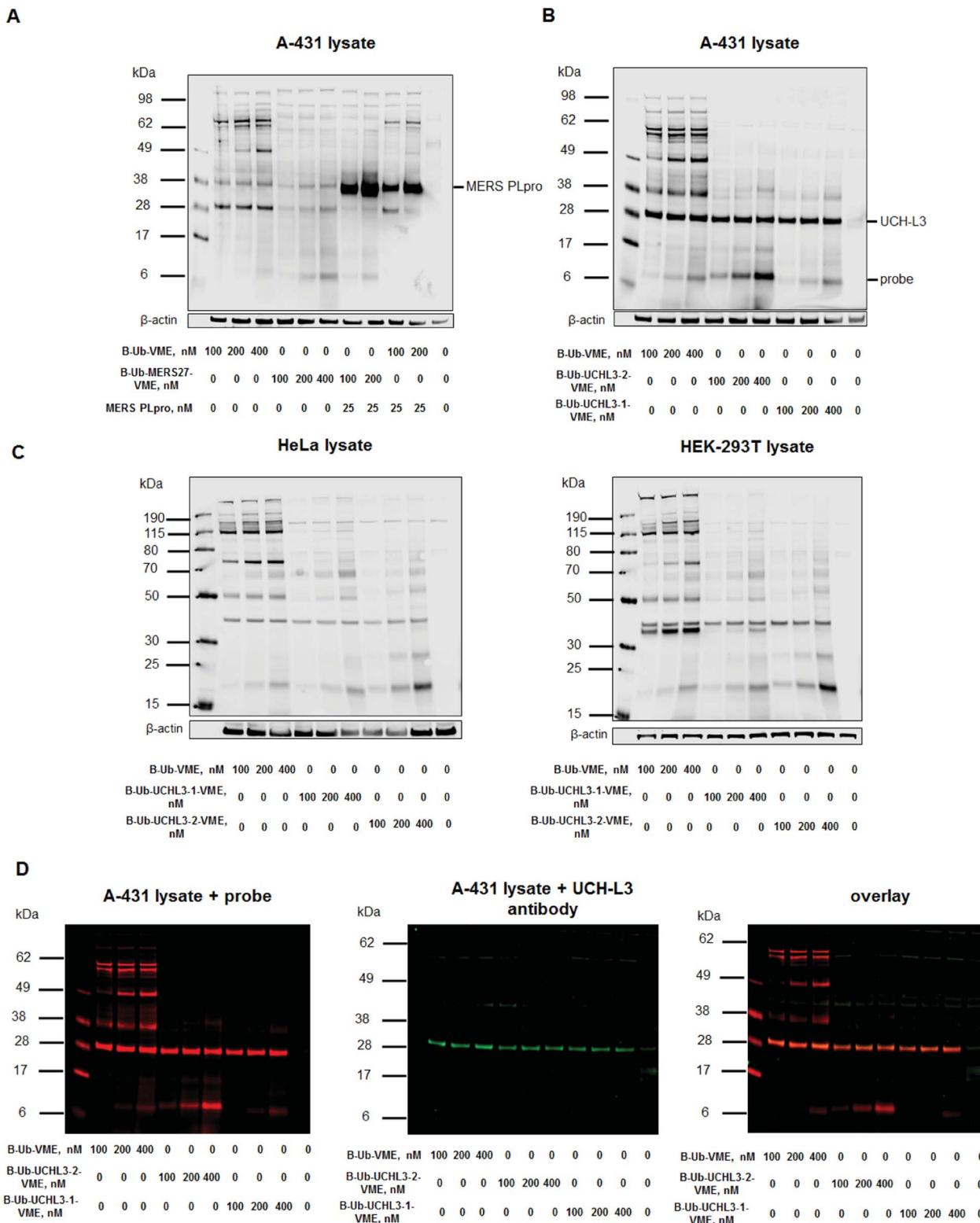


Fig. 5 Ub-based probe selectivity. DUB labeling in cell lysates by B-Ub-VME and B-Ub-MERS27-VME (A) and B-Ub-UCHL3-1-VME and B-Ub-UCHL3-2-VME (B and C). (D) Detection of UCH-L3 in A-431 cell lysates using (1) Ub-based probes (using a streptavidin Alexa Fluor 647 conjugate) and (2) a UCH-L3 antibody. A-431 cell lysate was incubated with three different probe concentrations (100, 200 and 400 nM) for 30 min at 37 °C. β-Actin was used as a loading control.

with glycine was attached to the Ub[1-75]-NH-NH<sub>2</sub> C-terminus, and the final product was purified and analyzed by HPLC and HRMS (Fig. 3).

Kinetic analysis of fluorescent Ub-based substrates revealed that by incorporation of unnatural amino acids in the Ub C-terminus, selective and active substrates for DUBs can be synthesized (Fig. 2F and G). The designed substrates were better hydrolyzed by MERS PLpro (Ub-MERS27-ACC) and UCH-L3 (Ub-UCHL3-1-ACC and Ub-UCHL3-2-ACC) than Ub-ACC, used as a reference substrate (Table 1). Ub-UCHL3-2-ACC was not recognized by MERS PLpro. Ub-MERS27-ACC was poorly hydrolyzed by UCH-L3 (almost 200 times weaker than Ub-ACC).

### 3.4. Design of selective Ub-based probes

One of the main aims of this work was to design Ub-based probes for highly selective detection of the investigated DUBs. Thus, we converted our Ub-based substrates to probes by incorporating a detection tag on the N-terminus and an irreversible reactive group on the C-terminus. VME was applied as an electrophilic warhead due to its broad reactivity toward DUBs.<sup>11</sup> The biotin tag was separated from the peptide sequence by using a 6-aminohexanoic acid linker (6-Ahx). All Ub-based probes were synthesized using the same synthetic strategy as that used for the Ub-based substrates (Fig. 3). The kinetic analysis with recombinant enzymes revealed that B-Ub-UCHL3-2-VME was a potent and selective probe for UCH-L3 (Fig. 4A). B-Ub-MERS27-VME was selective toward MERS PLpro (a much higher probe concentration was needed to inhibit UCH-L3). Furthermore, this probe was more potent toward MERS PLpro than the control probe B-Ub-VME (Fig. 4B). The results were generally consistent with the Ub-based substrate kinetic data; however, B-Ub-UCHL3-2-VME was a more potent probe for UCH-L3 than B-Ub-UCHL3-1-VME. We also determined  $k_{obs}/I$  value of B-Ub-VME ( $7660 \pm 475 \text{ M}^{-1} \text{ s}^{-1}$ ) and B-Ub-MERS27-VME ( $22\,144 \pm 1431 \text{ M}^{-1} \text{ s}^{-1}$ ) toward MERS PLpro, but we could not determine  $k_{obs}/I$  value of B-Ub-VME, B-Ub-UCHL3-1-VME and B-Ub-UCHL3-2-VME toward UCH-L3 due to the fact that both Ub based substrates are very efficiently interacting with binding site, thus making final data not really reliable.

### 3.5. Detection of DUBs in cell lysates

To assess Ub-based probe selectivity, we incubated cell lysates with various probe concentrations. Since the level of RNA expression coding UCH-L3 is high in the A-431 cell line, we selected it for model DUB labeling experiments.<sup>41</sup> Moreover, we have carried out labelling experiments in another two cell lines – HeLa and HEK-293T to check selectivity of UCH-L3 labelling with different subset of DUBs. Due to the lack of MERS PLpro expression in human cell lines, we added the recombinant enzyme to cell lysates during incubation with probes. The concentration of MERS PLpro used in cell lysate experiment was determined in another assay (Fig. S2†). These DUB assays in cell lysates with designed probes revealed that our probes (B-Ub-MERS27-VME, B-Ub-UCHL3-1-VME and B-Ub-UCHL3-2-VME) displayed high selectivity toward MERS PLpro and UCH-L3 compared to the reference probe that labeled many cellular

DUBs (Fig. 5A–C, S3A and C†). Labeling of UCH-L3 by our probes in A-431 cell lysate was confirmed by immunoblotting (Fig. 5D). To further demonstrate that UCH-L3 labelling is probe dependant, we have carried out pull-down experiments using B-Ub-VME, B-Ub-UCHL3-1-VME and B-Ub-UCHL3-2-VME on two cell lysates (A-431 and HeLa) (Fig. S3D in the ESI†).

## 4. Conclusions

DUBs constitute potential therapeutic targets in the treatment of viral diseases, neurodegenerative diseases and cancer.<sup>4,14,42</sup> The hottest example here is present COVID-19 outbreak caused by SARS-CoV-2. Current studies indicate that this coronavirus is similar to SARS and MERS. Provided by us, this technology could be used for fast delivery of a selective SARS-CoV-2 PLpro substrate and probe, which could effectively be used as a screen/reporter assay to find drugs that can specifically inhibit the SARS-CoV-2 PLpro DUB by screening infected patient cells without the need to purify DUB from infected patient cells. In classic examples in which DUBs are involved, enzymes belonging to the USP subclass play a key role in many biological processes and pathological conditions, including DNA repair, stabilization/degradation of the p53 protein, and progression of lung, breast, kidney and prostate cancer.<sup>43,44</sup> UCH-L3 is overexpressed in triple-negative breast cancer (TNBC), which correlates with poor prognosis.<sup>28</sup> This enzyme is also involved in the progression of pancreatic cancer, but little is known about its precise functions.<sup>29</sup> This is due to the lack of selective chemical tools for the study of this enzyme; such tools would enable the determination of the role of this enzyme in normal and pathological states and could offer leading structures and chemical starting points for drug development. DUBs are highly specific toward Ub; thus, the chemical tools most commonly used to investigate DUBs are based on a Ub scaffold. These broad-spectrum tools are valuable in DUB profiling assays. However, substrates and ABPs recognized by one or a narrow subset of DUBs would enable more precise examination of this group of enzymes. To date, Ub-based probe selectivity toward one or a narrow subset of DUBs has been achieved by (1) selecting various C-terminal reactive electrophilic groups;<sup>11,45</sup> (2) modifying the Ub scaffold length;<sup>46</sup> (3) synthesizing poly-Ub chains with different linkage topologies;<sup>16</sup> and (4) introducing various mutations in the Ub sequence.<sup>21,22</sup> Our previous work showed that Ub-based substrate selectivity can be accomplished by modification of the C-terminal LRGG motif of Ub.<sup>23</sup> The introduction of natural amino acids other than canonical LRG at the P4–P2 positions in the Ub derivative provided selectivity but reduced DUB activity by 20–50 times compared to the natural analog.

Based on this finding, we decided to incorporate unnatural amino acids in the C-terminal motif of Ub to improve DUB activity and selectivity toward Ub-based substrates and ABPs. For precise analysis of the DUB binding pocket architecture, we utilized a combinatorial approach and synthesized three tetrapeptide fluorogenic substrate libraries. Due to the low tetrapeptide substrate hydrolysis rate of DUBs, three positions in the combinatorial library (P4 and P3 sublibraries) were fixed, and



one contained an equimolar mixture of 19 amino acids. The P2 sublibrary was designed by fixing the P4, P3 and P1 positions as canonical amino acids present in Ub. The P2 position contained 128 natural and unnatural amino acids. This approach allowed us to achieve the highest possible substrate concentration in each position of the sublibrary during screening assays. We did not examine P1 substrate specificity because crystal structures of DUBs showed that this position can be occupied only by a glycine residue.<sup>13</sup> Library screening revealed that DUBs possess wide substrate specificity at the P4 and P3 positions. The DUBs recognized some unnatural amino acids even better than leucine and arginine at the P4–P3 positions. These results shed new light on DUB substrate preferences and the architecture of their binding sites. Kinetic analysis of tetrapeptide fluorogenic substrates revealed that the designed substrates with unnatural amino acids were selective and 2–3 times better recognized than the control substrate (Ac-LRGG-ACC); however, these substrates were still very poorly hydrolyzed. To improve DUB activity toward the designed substrates, we decided to synthesize unnatural Ub by incorporating the selected peptide sequences instead of the canonical LRGG sequence at the C-terminus and adding a fluorescent tag. DUBs possess two recognition regions that are required for effective Ub substrate hydrolysis.<sup>47</sup> The first region (being the secondary binding site or the exosite) interacts with the Ub surface, while the second region is the active center of the DUB, where the C-terminal LRGG motif is bound. Ub binding through the first region (distant from the active site cleft) leads to large conformational changes in DUB binding pockets that are required for effective catalysis of substrates.<sup>47</sup> Due to the nature of Ub binding by DUBs, we were expecting that translation from tetrapeptide substrates to full-length Ub sequences containing 2–3 unnatural amino acids on the C-terminus may result in a slight loss of selectivity. Despite the lack of hydrolysis of the MERS S27 tetrapeptide substrate by UCH-L3, kinetic analysis of the designed Ub-based substrates revealed that the MERS PLpro substrate (Ub-MERS27-ACC) was recognized by UCH-L3, but this recognition was almost 200 times weaker than that of the reference Ub-ACC. The UCH-L3 substrate (Ub-UCHL3-2-ACC) was not hydrolyzed by MERS PLpro even at high concentrations. Furthermore, these substrates were more efficiently hydrolyzed by DUBs than Ub-ACC. Interestingly, UCH-L3 S1, an approximately 3-fold weaker UCH-L3 tetrapeptide substrate than UCH-L3 S2, after conversion to the Ub-UCHL3-1-ACC substrate, had an almost 1.5-fold higher kinetic efficiency constant than the UCH-L3 S2 analog Ub-UCHL3-2-ACC. Nevertheless, both Ub-UCHL3-1-ACC and Ub-UCHL3-2-ACC displayed high selectivity toward UCH-L3 and exhibited kinetic parameters superior to those of Ub-ACC. Both of these features make these substrates very valuable alternatives to standard reagents used in studies of UCH-L3.

One of the most commonly used chemical tools in protease investigation are ABPs. Therefore, we converted our Ub-based fluorogenic substrates to Ub-based probes by replacing the ACC fluorophore with the VME electrophilic group and attaching a biotin tag to the N-terminus. Cell lysate assays confirmed the dramatically high selectivity of these engineered Ub-based

probes toward the investigated enzymes, especially when compared to the reference probe B-Ub-VME, which labeled many cellular DUBs. These results indicate that Ub-based selective chemical tools for DUBs can be obtained by introducing unnatural amino acids into the C-terminal Ub motif.

In summary, we demonstrated that our approach can be successfully applied in the design and synthesis of selective mono-Ub substrates and ABPs. Precise analysis of DUB binding pocket architecture using defined and combinatorial libraries with natural and unnatural amino acids allowed us to extract key residues that were introduced into the C-terminal motif of Ub and provided selectivity toward the investigated enzymes. Our findings expand the knowledge of DUBs, as well as the existing ‘toolbox’ of Ub-based biochemical tools to study this group of enzymes. Moreover, the presented chemical approach may be beneficial for the development of (a) new tools for the investigation of DUBs from distinct subclasses; (b) more selective tools based on poly-Ub chains with different topologies; and (c) new selective tools for studies of other Ub-like modifiers (UBLs), *e.g.*, SUMO proteins, ISG15 or Nedd8, and their conjugation/deconjugation biochemical machinery.

## Author contributions

M. D. conceptualized the study; M. D., W. R. and M. Z. designed the research; W. R. and M. Z. performed the research and collected data; S. J. S., M. B. and T. T. H. contributed enzymes; M. D., W. R. and M. Z. analyzed and interpreted the data and wrote the manuscript; and S. J. S., M. B. and T. T. H. critically revised the manuscript.

## Conflicts of interest

There are no conflicts to declare.

## Acknowledgements

This project was supported by the National Science Center grant 2015/17/N/ST5/03072 (Preludium 9) in Poland (W. R.) and the “TEAM/2017-4/32” project, which is carried out within the TEAM program of the Foundation for Polish Science, co-financed by the European Union under the European Regional Development Fund (M. D.). W. R. is a beneficiary of a START scholarship from the Foundation for Polish Science. NIH grants GM107257 (T. T. H.), ES025166 (T. T. H.) and F32GM100598 (M. B.) provided additional support for this work.

## References

- 1 Z. M. Eletr and K. D. Wilkinson, *Biochim. Biophys. Acta*, 2014, **1843**, 114–128.
- 2 A. Y. Amerik and M. Hochstrasser, *Biochim. Biophys. Acta*, 2004, **1695**, 189–207.
- 3 T. E. Mevissen, M. K. Hospenthal, P. P. Geurink, P. R. Elliott, M. Akutsu, N. Arnaudo, R. Ekkebus, Y. Kulathu, T. Wauer, F. El Oualid, S. M. Freund, H. Ovaa and D. Komander, *Cell*, 2013, **154**, 169–184.



4 A. M. Mielech, A. Kilianski, Y. M. Baez-Santos, A. D. Mesecar and S. C. Baker, *Virology*, 2014, **450**, 64–70.

5 H. A. Lindner, N. Fotouhi-Ardakani, V. Lytvyn, P. Lachance, T. Sulea and R. Menard, *J. Virol.*, 2005, **79**, 15199–15208.

6 Y. T. Li, Y. C. Huang, Y. Xu, M. Pan and Y. M. Li, *Tetrahedron*, 2016, **72**, 4085–4090.

7 L. C. Dang, F. D. Melandri and R. L. Stein, *Biochemistry*, 1998, **37**, 1868–1879.

8 U. Hassiepen, U. Eidhoff, G. Meder, J. F. Bulber, A. Hein, U. Bodendorf, E. Lorthiols and B. Martoglio, *Anal. Biochem.*, 2007, **371**, 201–207.

9 K. R. Love, A. Catic, C. Schlieker and H. L. Ploegh, *Nat. Chem. Biol.*, 2007, **3**, 697–705.

10 F. E. Reyes-Turcu, K. H. Ventii and K. D. Wilkinson, *Annu. Rev. Biochem.*, 2009, **78**, 363–397.

11 A. Borodovsky, H. Ovaa, N. Kolli, T. Gan-Erdene, K. D. Wilkinson, H. L. Ploegh and B. M. Kessler, *Chem. Biol.*, 2002, **9**, 1149–1159.

12 D. S. Hewings, J. A. Flygare, M. Bogyo and I. E. Wertz, *FEBS J.*, 2017, **284**, 1555–1576.

13 J. A. Ronau, J. F. Beckmann and M. Hochstrasser, *Cell Res.*, 2016, **26**, 441–456.

14 J. A. Harrigan, X. Jacq, N. M. Martin and S. P. Jackson, *Nat. Rev. Drug Discovery*, 2018, **17**, 57–78.

15 P. Gopinath, S. Ohayon, M. Nawatha and A. Brik, *Chem. Soc. Rev.*, 2016, **45**, 4171–4198.

16 A. Iphofer, A. Kummer, M. Nimtz, A. Ritter, T. Arnold, R. Frank, J. van den Heuvel, B. M. Kessler, L. Jansch and R. Franke, *Chembiochem*, 2012, **13**, 1416–1420.

17 G. R. Li, Q. Liang, P. Gong, A. H. Tencer and Z. H. Zhuang, *Chem. Commun.*, 2014, **50**, 216–218.

18 D. Flierman, G. J. V. van Noort, R. Ekkebus, P. P. Geurink, T. E. T. Mevissen, M. K. Hospenthal, D. Komander and H. Ovaa, *Cell Chem. Biol.*, 2016, **23**, 472–482.

19 M. Bokes, W. Rut, P. Kasperkiewicz, M. P. C. Mulder, H. Ovaa, M. Drag, C. D. Lima and T. T. Huang, *Biochem. J.*, 2015, **468**, 215–226.

20 M. Bokes, G. J. V. van Noort, R. Ekkebus, H. Ovaa, T. T. Huang and C. D. Lima, *Mol. Cell*, 2016, **62**, 572–585.

21 L. Gjonaj, A. Sapmaz, R. Gonzalez-Prieto, A. C. O. Vertegaal, D. Flierman and H. Ovaa, *Chem. Commun.*, 2019, **55**, 5075–5078.

22 L. Gjonaj, A. Sapmaz, D. Flierman, G. M. C. Janssen, P. A. van Veelen and H. Ovaa, *Chem. Sci.*, 2019, **10**, 10290–10296.

23 M. Drag, J. Mikolajczyk, M. Bokes, F. E. Reyes-Turcu, J. A. Ellman, K. D. Wilkinson and G. S. Salvesen, *Biochem. J.*, 2008, **415**, 367–375.

24 M. Poreba, R. Solberg, W. Rut, N. N. Lunde, P. Kasperkiewicz, S. J. Snipas, M. Mihelic, D. Turk, B. Turk, G. S. Salvesen and M. Drag, *Cell Chem. Biol.*, 2016, **23**, 1023–1035.

25 M. Poreba, W. Rut, M. Vizovisek, K. Groborz, P. Kasperkiewicz, D. Finlay, K. Vuori, D. Turk, B. Turk, G. S. Salvesen and M. Drag, *Chem. Sci.*, 2018, **9**, 2113–2129.

26 P. Kasperkiewicz, Y. Altman, M. D'Angelo, G. S. Salvesen and M. Drag, *J. Am. Chem. Soc.*, 2017, **139**, 10115–10125.

27 W. Rut, M. Poreba, P. Kasperkiewicz, S. J. Snipas and M. Drag, *J. Med. Chem.*, 2018, **61**, 5222–5234.

28 Z. W. Song, X. Y. Tu, Q. Zhou, J. Z. Huang, Y. P. Chen, J. Q. Liu, S. Lee, W. Kim, S. Nowsheen, K. T. Luo, J. Yuan and Z. K. Lou, *Cell Death Dis.*, 2019, **10**, 398.

29 Z. W. Song, J. H. Li, L. Zhang, J. Deng, Z. L. Fang, X. J. Xiang and J. P. Xiong, *Am. J. Cancer Res.*, 2019, **9**, 1970–1981.

30 S. Bondalapati, M. Jbara and A. Brik, *Nat. Chem.*, 2016, **8**, 407–418.

31 D. J. Maly, F. Leonetti, B. J. Backes, D. S. Dauber, J. L. Harris, C. S. Craik and J. A. Ellman, *J. Org. Chem.*, 2002, **67**, 910–915.

32 M. Poreba, G. S. Salvesen and M. Drag, *Nat. Protoc.*, 2017, **12**, 2189–2214.

33 M. Poreba, A. Szalek, P. Kasperkiewicz and M. Drag, *Methods Mol. Biol.*, 2014, **1133**, 41–59.

34 F. El Oualid, R. Merkx, R. Ekkebus, D. S. Hameed, J. J. Smit, A. de Jong, H. Hilkmann, T. K. Sixma and H. Ovaa, *Angew. Chem., Int. Ed. Engl.*, 2010, **49**, 10149–10153.

35 A. de Jong, R. Merkx, I. Berlin, B. Rodenko, R. H. Wijdeven, D. El Atmioui, Z. Yalcin, C. N. Robson, J. J. Neefjes and H. Ovaa, *Chembiochem*, 2012, **13**, 2251–2258.

36 J. Liang, G. M. Fang, X. L. Huang, Z. Q. Mei, J. Li, C. L. Tian and L. Liu, *Sci. China: Chem.*, 2013, **56**, 1301–1306.

37 Y. C. Huang, C. C. Chen, S. J. Li, S. Gao, J. Shi and Y. M. Li, *Tetrahedron*, 2014, **70**, 2951–2955.

38 J. S. Zheng, S. Tang, Y. K. Qi, Z. P. Wang and L. Liu, *Nat. Protoc.*, 2013, **8**, 2483–2495.

39 E. Vedrenne, H. Dupont, S. Oualef, L. Elkaim and L. Grimaud, *Synlett*, 2005, 670–672.

40 P. Kasperkiewicz, M. Poreba, S. J. Snipas, H. Parker, C. C. Winterbourn, G. S. Salvesen and M. Drag, *Proc. Natl. Acad. Sci. U. S. A.*, 2014, **111**, 2518–2523.

41 <https://www.proteinatlas.org/ENSG00000118939-UCHL3/> summary/rna.

42 P. D'Arcy, X. Wang and S. Linder, *Pharmacol. Ther.*, 2015, **147**, 32–54.

43 M. I. Davis and A. Simeonov, *Drug target review*, 2015, vol. 2, pp. 60–64.

44 M. J. Young, K. C. Hsu, T. E. Lin, W. C. Chang and J. J. Hung, *J. Biomed. Sci.*, 2019, **26**, 42.

45 K. R. Love, R. K. Pandya, E. Spooner and H. L. Ploegh, *ACS Chem. Biol.*, 2009, **4**, 275–287.

46 A. Borodovsky, H. Ovaa, W. J. Meester, E. S. Venanzi, M. S. Bogyo, B. G. Hekking, H. L. Ploegh, B. M. Kessler and H. S. Overkleeft, *Chembiochem*, 2005, **6**, 287–291.

47 T. A. Soboleva and R. T. Baker, *Curr. Protein Pept. Sci.*, 2004, **5**, 191–200.

