Generation of the UFM1 Toolkit for Profiling UFM1-Specific Proteases and Ligases

Katharina F. Witting, Gerbrand J. van der Heden van Noort, Christian Kofoed, Cami Talavera Ormeño, Dris el Atmioui, Monique P. C. Mulder,* and Huib Ovaa*

Abstract: Ubiquitin-fold modifier 1 (UFM1) is a reversible post-translational modifier that is covalently attached to target proteins through an enzymatic cascade and removed by designated proteases. Abnormalities in this process, referred to as Ufmylation, have been associated with a variety of human diseases. Given this, the UFM1-specific enzymes represent potential therapeutic targets; however, understanding of their biological function has been hampered by the lack of chemical tools for activity profiling. To address this unmet need, a diversifiable platform for UFM1 activity-based probes (ABPs) utilizing a native chemical ligation (NCL) strategy was developed, enabling the generation of a variety of tools to profile both UFM1 conjugating and deconjugating enzymes. The use of the probes is demonstrated in vitro and in vivo for monitoring UFM1 enzyme reactivity, opening new research avenues.

Post-translational modification (PTM) of proteins by chemical groups, peptides, complex molecules, or even small proteins facilitates dynamic protein diversification to modulate cellular responses. Ubiquitination is one of the most common PTMs and a number of Ub-like proteins (Ubls) have been subsequently identified. Ubiquitin-fold modifier 1 (UFM1) is one of the recently identified Ubls and displays a similar tertiary structure, yet it has little sequence identity to Ubiquitin (Ub). Analogous to ubiquitin, it is covalently attached to the lysine residues of its substrates by the sequential action of three dedicated enzymes-E1 (UBA5), E2 (Ufc1), and E3 (Ufl1) and is cleaved by UFM1-specific proteases (Ufsps).[3] This process, referred to as Ufmylation, is initiated by the adenylation of the exposed C-terminal glycine of mature UFM1 and subsequent nucleophilic reaction with the active-site cysteine of UBA5. The resulting high-energy thioester bond allows the transfer onto the catalytic site cysteine of the E2 enzyme Ufc1, in a trans-thioesterification reaction. Lastly, the E3-like enzyme Ufl1 mediates the transfer of activated UFM1 onto the lysine residues of the protein substrates resulting in the formation of an isopeptide linkage. In addition to releasing UFM1 from its substrates, the UFM1 specific proteases-Ufsp1 and Ufsp2mediate the maturation of pro-UFM1.[5] Although Ufmylation has been connected to biological processes including ER homeostasis,[4,6] vesicle trafficking,[5] blood progenitor development and differentiation,[7,8] G-coupled protein receptor (GPCR) maturation,[9] transcriptional control,[10] mitosis,[11] and autophagy,[7,12] the underlying mechanisms remain to be studied. Furthermore, abnormalities in the UFM1 cascade are reported to be associated with a number of human diseases, including cancer,[13] diabetes,[14] schizophrenia,[15] and ischemic heart disease[6] and to play a pivotal role in embryonic development and hematopoiesis.[8,16] Notwithstanding the biochemical and structural studies of UFM1-conjugating and deconjugating enzymes that have been undertaken,[7,21] their biological function remains enigmatic primarily owing to the lack of activity-based reagents. By contrast, diverse reagents and ABPs have been developed for both Ub-conjugating and deconjugating enzymes,[22–27] and have been expanded to Ubls such as SUMO[28–30] and Nedd8.[24,26] This advancement of assay and activity-based reagents has greatly propelled discoveries in the ubiquitin field, yet such a diversifiable synthetic platform for UFM1 needs to be developed.[31] While UFM1 has been prepared using multiple segment ligations based on KAH chemistry, this is a time-consuming process requiring the incorporation of a (S)-5-oxaproline building block at multiple sites.[32] We first attempted to generate UFM1 using linear solid-phase peptide synthesis (SPPS) by incorporating aggregation breakers such as pseudoproline[33] at permissible sites (Supporting Information, Figure S1). Although this linear synthesis approach yields full-length UFM1, which has been utilized in a recent study,[34] the synthesis wasn’t efficient, presumably owing to inefficient coupling of amino acid 36 onwards (see Figure S2). To circumvent this issue, we present herein a more practical two-segment native chemical ligation (NCL) approach[35,36] towards full-length UFM1 and UFM1 activity-based reagents (Figure 1). Given the increasing knowl-
edge on the importance of UFM1, our synthetic strategy gives access to valuable tools that allow the in vitro and in vivo characterization of enzymatic activity, thereby enabling insights into the dynamics of Ufmylation.

To improve the UFM1 synthesis, we devised a practical native chemical ligation (NCL) strategy to generate full-length UFM1 based on a N-terminal peptide thioester fragment (AA 1–44) and C-terminal peptide fragment (AA 45–83) with alanine at position 45 replaced by cysteine (Scheme 1). This method permitted the generation of full-length UFM1 and a complete repertoire of probes in an efficient manner using a minimal amount of building blocks (Scheme 1).

Using SPPS and standard coupling conditions (namely 4 equiv Fmoc-protected amino acid, 4 equiv PyBOP, 8 equiv DIPEA, and double coupling cycles), we prepared the N-terminal fragment (AA 1–44) on a hyper-acid labile trityl resin to give amine 1 or functionalized its N-termini with a Rho-amine 110 fluorophore (2).[17] Then, selective cleavage from the resin with 20% hexafluoro-2-propanol (HFIP) liberated the C-terminal carboxylate leaving all other protective groups in place. The C-terminal carboxylate was then converted into the mercaptomethylpropionate thioester by reaction with mercaptomethylpropionate, followed by global deprotection and HPLC purification. Subsequently, the C-terminal fragment (A45C-83, 3) was obtained by Fmoc-based SPPS, followed by global deprotection and HPLC purification.

Native chemical ligation of the N-terminal thioester-fragment (1 or 2) and C-terminal cysteine-fragment (3) was performed under denaturing conditions in 8 M Gdn·HCl containing 100 mM tris(2-carboxyethyl)phosphine (TCEP) and 100 mM mercaptoethanesulfonic acid (MPAA) at pH 7.6 and 37°C. Conversion to the full-length UFM1 product (6 or 7) proceeded very efficiently in just 30 min. Precipitation from water and subsequent redissolution in Gdn·HCl followed by radical desulfurization using VA-044, glutathione, and TCEP for 16 h at 37°C resulted in formation of the full-length native UFM1 as the cysteine residue at the ligation site was converted into the native alanine. Subsequent HPLC purification followed by gel-filtration yielded UFM1 and Rho-UFM1 in 85% and 79% yield, respectively.

Correct folding of purified synthetic UFM1 (Figure 2A) was verified by circular dichroism (CD) spectroscopy (Figure 2B). To further verify the correct folding and thus biochemical function, we compared synthetic and recombinant UFM1 in an enzymatic reaction with UBA5 (E1). Efficient formation of the UBA5–UFM1 thioester proved that synthetic UFM1 is processed with the same efficiency as recombinant UFM1 (Figure 2C and Figure S3).

Having developed a convenient UFM1 synthesis, we then focused on the generation of a UFM1-based toolbox to enable...
the profiling and visualization of all enzymes involved in the Ufmylation cascade (Figure 1, Scheme 1). For this purpose, we synthesized UFM1-Dha to target the conjugating class of enzymes[26] and UFM1-PA to target the proteases.[24] To achieve this, we synthesized a C-terminal fragment lacking the last amino acid (A45C-82). Upon cleavage from the resin with HFIP, the C-termini were equipped with either Cys(Bn)-OMe (4) or propargylamide (5). After global deprotection and HPLC purification, the desired C-terminal fragments were obtained and subsequently used in a NCL reaction as described above. ABP precursors for the UFM1 ligase probe (8) was successfully constructed by NCL followed by a radical desulfurization reaction to obtain UFM1-Cys(Bn)-OMe. This was subsequently transformed by oxidative elimination with O-mesitylenesulfonylhydroxyl-amine (MSH). Finally, the methyl ester was hydrolysed, to generate the UFM1-Dha probe (8). ABPs targeting UFM1-specific cysteine proteases (9–10) were easily constructed by native chemical ligation as well. However, one limitation we encountered was that the remaining cysteine after NCL could not be desulphurized, as radical desulfurization conditions compromise the integrity of the propargyl warhead. But since the ligation position (AA 45) is situated outside the critical C-terminal region for UFM1 recognition by Ufsp enzymes,[38] we expect this small thiol group to be tolerated. Furthermore, the routine addition of 1,4-dithiothreitol (DTT) to the enzymatic assay reaction buffers reduces any potential disulfide bridges occurring at this cysteine that might affect enzymatic processing.

With these ABPs in hand, we tested their reactivity towards UFM1 ligases and proteases in cell lysates. Firstly, we assessed the reactivity of the E1 enzyme-UBA5 towards Rho-UFM1-Dha (8). Similar to UbDha, UFM1-Dha forms an adenylate with the C-terminal glycine of UFM1. The activated methylene group of this adenylate intermediate can then undergo a nucleophilic attack of the active-site cysteine to yield either the covalent UBA5-UFM1-thioether adduct or the UBA5-UFM1–thioester[39] (Figure 3A). Upon reacting UBA5 with the ligase probe Rho-UFM1-Dha (8), a complex is formed that remains stable in the presence of DTT, indicative of the thioether bond formed between the active site cysteine and the dehydroalanine moiety of 8 in contrast to the thioester bond formed between UBA5 and UFM1 (7) (Figure 3B and Figure S3). Furthermore, UFM1-Dha does not display cross-reactivity towards UBE1, the Ub-activating E1 enzyme, or towards the Ufm1-specific protease Ufsp1 underscoring the selectivity of this ABP (Figures S4 and S5).

Next, we tested our UFM1-PA protease probe.[34] As expected, Rho-UFM1-PA (10) generated by NCL (containing Cys-45) showed comparable reactivity towards Ufsp1 as the Rho-UFM1-PA probe generated by linear SPPS (containing the native Ala-45), indicating that the cysteine at this position does not interfere with protease recognition, which is in agreement with literature[38] (Figure S6). Given that human Ufsp1 is thought to be catalytically inactive (Figure S7), we assessed the reactivity of the active murine Ufsp1 instead. To this end, HEK293 cells transiently overexpressing Flag-tagged murine Ufsp1 or untransfected cells were lysed and incubated with Rho-UFM1-PA at 37°C (Figure 4A). Visualisation by in-gel fluorescence scanning followed by immunoblotting revealed that Ufsp1 engages more readily than Ufsp2 with this ABP[35] (Figure 4B and Figure S8). While murine Ufsp1 completely reacts with UFM1-PA after 45 min, endogenous Ufsp2 engages with UFM1-PA after prolonged incubation of nearly 6 h (Figure 4A,B). The two UFM1-specific proteases known to date are cysteine-proteases exhibiting a conserved papain-like fold with the classical catalytic Cys–His–Asp triad but share no sequence homology with DUBS

**Figure 3.** Reactivity of E1 (UBA5) towards UFM1. A) Scheme depicting the reactivity of UBA5 towards UFM1-Dha permitting the formation of either the thioether (a) or the thioester adduct (b). B) SDS-PAGE gel of reaction of synthetic UFM1 (6) or Rho-UFM1-Dha (8) and recombinant UBA5.

**Figure 4.** Assessment of UFM1-PA reactivity and specificity against a panel of cysteine protease subfamilies. A) Immunoblot showing time-dependent reactivity of Flag-Ufsp1 (murine) towards Rho-UFM1-PA (10), B) and time-dependent labelling of endogenous Ufsp2 (human). C,F) DUBs Flag-UCHL-1 and GFP-OTUB2 do not react with UFM1-PA. D,E) Flag-SENP6 and Flag-SENP1 react with S2-PA (=SUMO2-PA) but are unreactive towards UFM1-PA. For fluorescence scans and actin blots see Figure S8.
or Ub-like-protein-specific proteases (ULPs) but rather constitute a new cysteine protease subfamily.\textsuperscript{[20,21]} To assess whether deubiquitinating enzymes (DUBs) and SUMO-specific proteases (SENPs) display cross reactivity towards UFM1-PA, a panel of representative deubiquitinating and deSUMOylating enzymes were profiled. As expected, only UFM1 specific proteases recognize and bind specifically to UFM1-PA (Figure 4 C–F and Figure S8).

Subsequently, we addressed the efficacy of UFM1-PA by monitoring UFM1 protease activity in a cellular environment by electroporation of Rho-UFM1-PA (10) into HeLa cells ectopically expressing Flag-Ufsp1 or its catalytically inactive C53A mutant as previously described.\textsuperscript{[26]} In-gel fluorescence analysis followed by immunoblotting revealed labeling of endogenous Ufsp2 in cells with Rho-UFM1-PA, while in Ufsp2-depleted HeLa cells reactivity was abolished (see Figure 5A and Figures S9 and S10). As expected, only active Flag-Ufsp1 reacted with UFM1-PA but not its catalytically inactive version. Having demonstrated that our UFM1 ABP probe can be used to monitor enzymatic reactivity in living cells, we used the fluorescent ABP to visualize Ufsp reactivity by confocal microscopy. After introduction of Rho-UFM1-PA by electroporation into either unmodified or HeLa cells ectopically transfected with murine Flag-Ufsp1, probe distribution was observed both throughout the cytoplasm and the nucleus (Figure 5B). However, incorporation of 10 into HeLa cells ectopically expressing murine Flag-Ufsp1 or its catalytically inactive C53A mutant, showed substantial co-localization with wild-type but not with the catalytically incompetent enzyme (Figure 5C).

In conclusion, we present a practical native chemical ligation-based synthetic approach for generating UFM1 activity-based probes. Using this facile strategy, a wide variety of UFM1 ABPs equipped with fluorescent tags, as well as diverse reactive groups targeting proteases or ligases are accessible. The strategies presented in this study highlight a variety of assays and possibilities to utilize our ABPs in the interrogation of enzymatic activities in the Ufmylation cascade. The straightforward nature of the experimental setup is expected to make them readily adoptable to comparative profiling in the presence of various perturbations, facilitating the understanding of UFM1 biology and additionally potentiate the discovery of specific UFM1 enzyme inhibitors. We anticipate that our toolset will greatly facilitate ongoing studies on the UFM1 enzyme cascade.

Acknowledgements

We would like to thank Dr. Aimee Boyle at Leiden University for assistance with the CD measurements. This work was supported by a VICI grant (NWO-724.013.002) to H.O. from the Netherlands Organization for Scientific Research (N.W.O.).

Conflict of interest

The authors declare no conflict of interest.

Keywords: activity-based probes · chemical biology · native chemical ligation · post-translational modifications · UFM1

How to cite: Angew. Chem. Int. Ed. 2018, 57, 14164–14168

Angew. Chem. 2018, 130, 14360–14364

[1] M. Komatsu, T. Chiba, K. Tatsumi, S. Iemura, I. Tanida, N. Okazaki, T. Ueno, E. Kominami, T. Natsume, K. Tanaka, EMBO J. 2004, 23, 1977–1986.
[2] Y. Wei, X. Xu, Genomics Proteomics Bioinf. 2016, 14, 140–146.
[3] S. H. Kang, G. R. Kim, M. Seong, S. H. Baek, J. H. Scol, O. S. Bang, H. Ovaa, K. Tatsumi, M. Komatsu, K. Tanaka, C. H. Chung, J. Biol. Chem. 2007, 282, 5256–5262.
[4] J. Liu, Y. Wang, L. Song, L. Zeng, W. Yi, T. Liu, H. Chen, M. Wang, Z. Ju, Y. S. Cong, Nat. Commun. 2017, 8, 14166.
[5] Y. Zhang, M. Zhang, J. Wu, G. Lei, H. Li, PLoS One 2012, 7, e48587.
[6] A. Azfer, J. Niu, L. M. Rogers, F. M. Adamski, P. E. Kolattukudy, Am. J. Physiol. Heart Circ. Physiol. 2006, 291, H1411–1420.
[7] M. Zhang, X. Zhu, Y. Zhang, Y. Cai, J. Chen, S. Sivaprakasam, A. Gurav, W. Pi, L. Makala, J. Wu, B. Pace, D. Tuan-Lo, V. Ganapathy, N. Singh, H. Li, Cell Death Differ. 2015, 22, 1922–1934.
[8] Y. Cai, W. Pi, S. Sivaprakasam, X. Zhu, M. Zhang, J. Chen, L. Makala, C. Lu, J. Wu, Y. Teng, B. Pace, D. Tuan, N. Singh, H. Li, PLoS Genet. 2015, 11, e1005643.
[9] C. Chen, E. Itakura, K. P. Weber, R. S. Hegde, M. de Bono, PLoS Genet. 2014, 10, e1004682.
[10] X. Hu, Q. Pang, Q. Shen, H. Liu, J. He, J. Wang, J. Xiong, H. Zhang, F. Chen, Int. J. Mol. Med. 2014, 33, 1539–1546.
[1] Y. Merbl, P. Refour, H. Patel, M. Springer, M. W. Kirschner, Cell 2013, 152, 1160–1172.

[2] R. DeJesus, F. Moretti, G. McAllister, Z. Wang, P. Bergman, S. Liu, E. Frias, J. Alford, J. S. Reicce-Hoyes, A. Lindeman, J. Kellilher, C. Russ, J. Knehr, W. Carbone, M. Beibel, G. Roma, A. Ng, J. A. Tallarico, J. A. Porter, R. J. Xavier, C. Mickanin, L. O. Murphy, G. R. Hoffman, B. Nfuleler, eLife 2016, 5, e17290.

[3] H. M. Yoo, S. H. Kang, J. Y. Kim, J. E. Lee, M. W. Seong, S. W. Lee, S. H. Ka, Y. S. Sou, M. Komatsu, K. Tanaka, S. T. Lee, D. Y. Noh, S. H. Baeck, Y. J. Jeon, C. H. Chung, Mol. Cell 2014, 56, 261–274.

[4] H. Lu, Y. Yang, E. M. Allister, N. Wijesekara, M. B. Wheeler, Mol. Cell Proteomics 2008, 7, 1434–1451.

[5] M. D. Rubio, K. Wood, V. Haroutumian, J. H. Meador-Woodruff, NeuroPsychopharmacology 2013, 38, 1910–1920.

[6] Y. Cai, N. Singh, H. Li, Exp. Hematol. 2016, 44, 442–446.

[7] W. Oweis, P. Padala, F. Hassouna, E. Cohen-Kfir, D. R. Gibbs, E. A. Todd, C. E. Berndsen, R. Wiener, Cell Rep. 2016, 16, 3113–3120.

[8] S. Habisov, J. Huber, Y. Ichimura, M. Akutsu, N. Rogova, F. Loehr, D. G. McEwan, T. Johansen, I. Dikie, V. Doetsch, M. Komatsu, V. V. Rogov, V. Kirkin, J. Biol. Chem. 2016, 291, 9025–9041.

[9] T. Mizushima, K. Tatsumi, Y. Ozaki, T. Kawakami, A. Suzuki, K. Ogashara, M. Komatsu, E. Kominami, K. Tanaka, T. Yamane, Biochem. Biophys. Res. Commun. 2007, 362, 1079–1084.

[10] B. H. Ha, H. C. Ahn, S. H. Kang, K. Tanaka, C. H. Chung, E. E. Kim, J. Biol. Chem. 2008, 283, 14893–14900.

[11] B. H. Ha, Y. J. Jeon, S. C. Shin, K. Tatsumi, M. Komatsu, K. Tanaka, C. M. Watson, G. Wallis, C. H. Chung, E. E. Kim, J. Biol. Chem. 2011, 286, 10248–10257.

[12] M. P. C. Mulder, F. El Ouailid, J. ter Beek, H. Ovaa, ChemBioChem 2014, 15, 946–949.

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

[14] R. Ekkebus, S. I. van Kasteren, Y. Kulahtu, A. Scholten, I. Berlin, P. P. Geurink, A. de Jong, S. Goerdalayil, J. Neefjes, A. J. Heck, D. Komander, H. Ovaa, J. Am. Chem. Soc. 2013, 135, 2867–2870.

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

[16] M. P. C. Mulder, K. Witting, I. Berlin, J. N. Pruneda, K. P. Wu, J. G. Chang, R. Merkx, J. Bialas, M. Groettrup, A. C. Vertegaal, B. A. Schulman, D. Komander, J. Neefjes, P. El Ouailid, H. Ovaa, Nat. Chem. Biol. 2016, 12, 523–530.

[17] A. de Jong, K. Witting, R. Kooji, D. Flierman, H. Ovaa, Angew. Chem. Int. Ed. 2017, 56, 12967–12970; Angew. Chem. 2017, 129, 13147–13150.

[18] S. Sommer, N. D. Weikart, U. Linne, H. D. Mootz, Bioorg. Med. Chem. 2013, 21, 2511–2517.

[19] V. E. Albrow, E. L. Ponder, D. Fasci, M. Becke, G. S. Salvesen, M. Bogoy, Chem. Biol. 2011, 18, 722–732.

[20] M. P. C. Mulder, R. Merkx, K. Witting, D. Hamed, D. E. Atmioui, L. Lelieveld, F. Liebelt, J. Neefjes, I. Berlin, A. Vertegaal, H. Ovaa, Angew. Chem. Int. Ed. 2018, 57, 8958–8962; Angew. Chem. 2018, 130, 9096–9100.

[21] K. F. Witting, M. P. C. Mulder, H. Ovaa, J. Mol. Biol. 2017, 429, 3388–3394.

[22] A. O. Ogunkoya, V. R. Pattabiraman, J. W. Bode, Angew. Chem. Int. Ed. 2012, 51, 9693–9697; Angew. Chem. 2012, 124, 9831–9835.

[23] O. Melnyk, J. Vicogne, Tetrahedron Lett. 2016, 57, 4316–4324.

[24] T. Hermans, C. Pichilo, I. Woiwode, K. Klopfleisch, K. F. Witting, H. Ovaa, U. Baumann, K. Hofmann, Nat. Commun. 2018, 9, 799.

[25] P. E. Dawson, T. W. Muir, I. Clark-Lewis, S. B. Kent, Science 1994, 266, 776–779.

[26] S. Kent, Bioorg. Med. Chem. 2017, 25, 4926–4937.

[27] P. P. Geurink, B. D. van Tol, D. van Dalen, P. J. Brandell, T. E. Mevissen, J. N. Pruneda, P. R. Elliott, G. B. van Tilburg, D. Komander, H. Ovaa, ChemBioChem 2016, 17, 816–820.

[28] K. H. Kim, B. H. Ha, E. E. Kim, FEBS Lett. 2018, 592, 263–273.

Manuscript received: August 10, 2018
Revised manuscript received: September 4, 2018
Accepted manuscript online: September 6, 2018
Version of record online: October 2, 2018