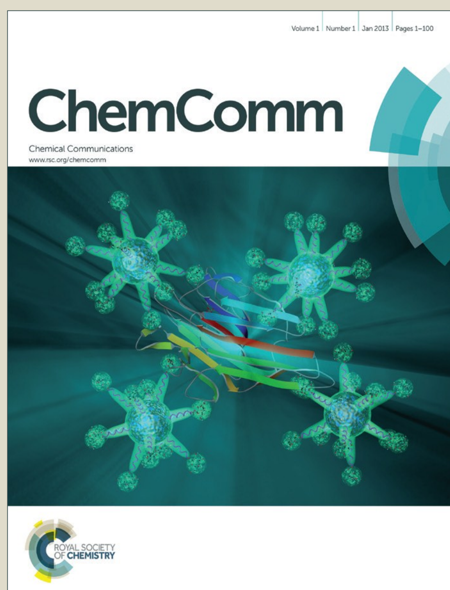
ChemComm

Accepted Manuscript



This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the Information for Authors.

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard Terms & Conditions and the Ethical guidelines still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/chemcomm

Please do not adjust margins

Journal Name

COMMUNICATION

Facile synthesis of covalent probes to capture enzymatic intermediates during E1 enzyme catalysis†

Heeseon An and Alexander V. Statsyuk*

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

www.rsc.org/

We report a facile synthetic strategy to prepare UBL-AMP electrophilic probes that form a covalent bond with the catalytic cysteine of cognate E1s, mimicking the tetrahedral intermediate of E1-UBL-AMP complex. These probes enable the structural and biochemical study of both canonical- and non-canonical E1s.

Ubiquitin and approximately ~20 other ubiquitin like proteins (UBL) regulate a vast array of cellular processes by reversibly modifying their substrates.[1] Each UBL system utilizes enzymatic cascades of cognate E1, E2, and E3 enzymes for the substrate conjugation, while isopeptidases remove UBLs from the substrates. E1 enzymes initiate the UBL conjugation process using a conserved catalytic mechanism.[2] Initially, a UBL and ATP bind to E1 enzymes, forming UBL-AMP complex and releasing PP_i (Fig. 1A). Subsequently, the catalytic cysteine of E1 attacks the reactive acyl-phosphate in UBL adenylate, resulting in the formation of E1~UBL thioester and AMP.

There are eight E1s known in humans, which are categorized as either canonical E1s or non-canonical E1s (Fig. 1B).[3] The canonical E1s have their catalytic cysteine ~30 Å away from the α-phosphate of ATP, thus large structural changes are required to form E1~UBL thioester.[4-6] In comparison, non-canonical E1s have their catalytic cysteine in closer proximity to the ATP binding site.

Despite the accumulated structural information on E1 enzymes, many questions regarding their structural and biochemical properties remain unanswered. For example, the catalytic cysteine of ATG7 is on a flexible loop and is ~7 Å away from the C-terminus of ATG8 (UBL), pointing in the opposite direction of the ATP binding centre.[7] How the catalytic cysteine rearranges for the nucleophilic attack on the C-terminal ATG8 adenylate is still unknown. Another example is UBA5, the least characterized of all E1s. The crystal structure of ATP-bound UBA5 shows that the catalytic cysteine is positioned on a long alpha helix located ~17 Å from the C-terminal Gly of Ufm1(UBL).[8] This structural feature of UBA5 raises interesting questions: how does UBA5 bring the cysteine close to its ATP binding centre for thioester formation? Does this require structural remodelling of the long alpha helix? Which amino acid residues are involved in the induction of such rearrangement?

Covalent trapping of ternary E1-UBL-AMP complexes is a useful approach that provides critical insights on the macromolecular structure and conformation of E1 enzymes during the catalytic cycle. To address this challenge, Olsen et al. reported an electrophilic SUMO-AVSN probe that could capture the remarkable conformational rearrangements of SUMO E1 during the catalytic cycle.[9] SUMO-AVSN mimics the structure of SUMO-AMP and contains a vinyl sulfonamide that covalently traps the catalytic cysteine of SUMO E1. As a result, it forms a covalently crosslinked ternary complex that resembles the tetrahedral intermediate of the native SUMO E1-SUMO-AMP complex.

Although the pioneering work on the SUMO-AVSN paves the path forward towards the creation of other UBL-AVSN probes, the following disadvantages limit broad application of UBL-AVSN probes to study other E1s[10]: (1) Over 10 synthetic steps for preparation of the precursor small molecule, (2) the last three residues of the UBL C-terminus must be CGG, which is significantly different from the native sequences of many UBLs, and (3) requirement of an intein based method to prepare engineered UBLs.[10] Given these limitations, an efficient approach for the synthesis of UBL-AMP probes bearing an electrophile is needed.

Here we report a facile synthetic strategy to prepare UBL-AMP mimic probes that can selectively label the catalytic cysteine of their cognate E1s (Fig. 1C). Our strategy relies on two key steps: native chemical ligation[11] and dehydroalanine chemistry.[12, 13] This approach has the following advantages: (1) the precursor AMP mimic for native chemical ligation (Fig 1C) can be prepared in 3 synthetic-steps[14], (2) electrophilic UBL-AMP probes preserve the natural C-terminal amino acid sequences of UBLs, (3) C-terminally activated UBL thioesters can be prepared either by using intein-based methods[15] or by E1-mediated chemoenzymatic synthesis,[16, 17] and (4) an alkyne tag on the adenine moiety provides an additional detection handle that can be used to detect and quantify the amount of E1-UBL-AMP complex using click chemistry.

Department of Chemistry, Center for Molecular Innovation and Drug Discovery, Chemistry of Life Processes Institute, Northwestern University, Silverman Hall, 2145 Sheridan Road, Evanston, Illinois 60208.

† Electronic Supplementary Information (ESI) available: See DOI: 10.1039/x0xx00000x

Please do not adjust margins

ChemComm Accepted Manuscript

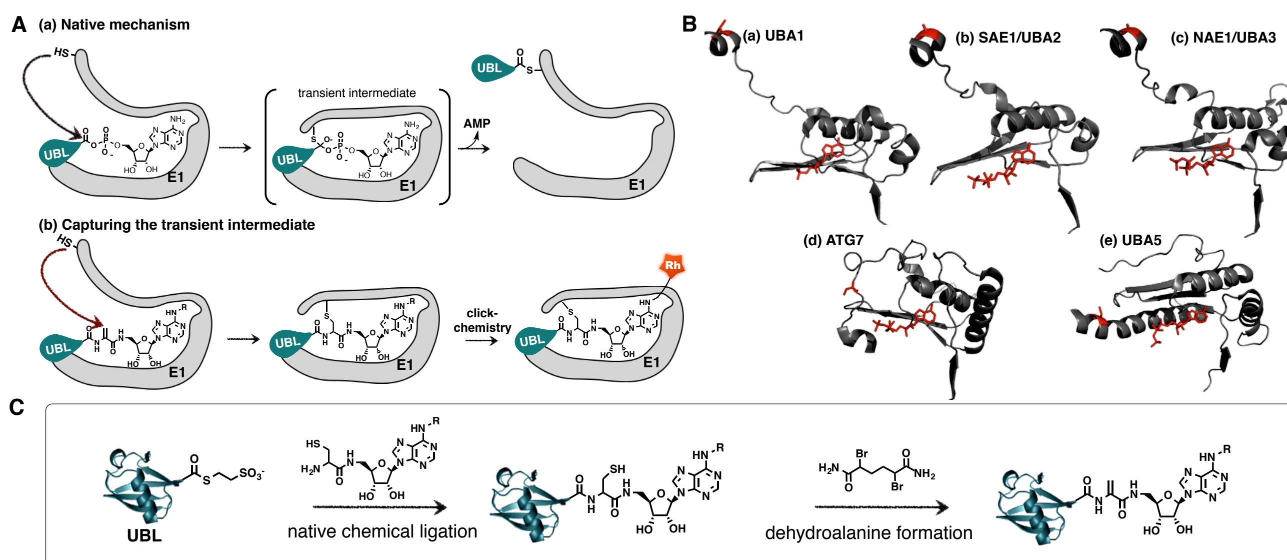


Fig. 1 (A) a: native mechanism of E1-UBL thioester formation b: a strategy to covalently trap tetrahedral E1-UBL-AMP reactive intermediate with UBL-AMP mimic probe. R: a hydrophobic substituent containing an alkyne functionality. Rh: Rhodamine dye (B) Crystal structures of ATP-bound E1 active centre. The catalytic cysteine and ATP are highlighted in red. (a)-(c): canonical E1s, (d)-(e): non-canonical E1s (C) Proposed synthesis of the protein probes.

Initially, we rationalized that UBL-electrophile probes lacking an adenine moiety will act as control probes to investigate the role of the adenine moiety during the E1~UBL thioester formation (Fig. 2A). We also hypothesized that we could modulate the binding affinity and labelling efficiency of our probes toward various E1s by installing different hydrophobic substituents at the N^6-position of the adenine moiety. Accordingly, we synthesized Pro-1, Pro-2, and Pro-3 precursor molecules (Figure 2A). We then developed a protocol for native chemical ligation and dehydroalanine chemistry to prepare electrophilic Ub-Probes for our initial model studies. Thus, C-terminally activated Flag-tagged ubiquitin (Ub~Mes) was prepared using intein methods.[15] Flag-Ub~Mes was then incubated with Pro 1-3 for 2 hours at room temperature with shaking. The formation of the native chemical ligation reaction products (i.e. Ub-Pro1-SH, Ub-Pro2-SH, and Ub-Pro3-SH) was confirmed by LC-MS analysis of the reaction mixture (Fig. 2A). Subsequently, the thiol functionality on the ubiquitin probes was converted to dehydroalanine using 2,5-dibromohexanediamide (hereafter di-bromide reagent).[12, 13] After dialysis, the formation of the final dehydroalanine probes was confirmed by LC-MS (Fig. 2B).

Following the successful synthesis of the electrophilic Ub-Probes, we asked if our strategy could be used to prepare other electrophilic UBL-probes. Specifically, we focused on ATG7-targeting UBL-probes. Activation of ATG8 (UBL) by ATG7 (E1) is critical for the formation of autophagosomes during autophagy. Therefore, small molecule modulators of ATG7 will serve as valuable tools to study autophagy and to validate ATG7 as a drug target.[18, 19] Investigating the structural remodelling of ATG7 during the catalytic cycle will lay the foundation for design of mechanism-based small molecule inhibitors of ATG7.

In humans, six known ATG8 homologs can be divided into three subfamilies: LC3 (3 genes), GABARAP (2 genes), and GATE-16.[20] For our purposes, we prepared C-terminally activated LC3~Mes using chemoenzymatic protocol.[17] The resulting LC3~Mes was incubated with Pro-1, Pro-2, or Pro-3 for 2 hours, followed by dehydrolalanine formation reaction to prepare LC3-Probes as confirmed by LC-MS analysis (Fig. 2C). Interestingly, we also observed that 15~30 % of LC3~Mes was hydrolyzed during the native chemical ligation reaction. This can be because the cysteine containing precursor small molecules (Pro 1-3) are less efficient in attacking the LC3~Mes thioester for the transthiolation reaction. According to the previous structural studies, ubiquitin has an extended C-terminal tail (RLR**GG**) exposed to solvent. However, Phe[119] in the C-terminus of LC3 (QET**FG**) has been suggested to contact the hydrophobic groove on LC3, forming compact "closed" conformation.[21] In summary, the developed synthetic approach would be generally applicable for the synthesis of a broad range of electrophilic UBL-AMP mimic probes.

Next, we investigated the mechanism-based labelling of E1 enzymes using electrophilic UBL probes (Fig. 3). Initially, 1 μM of wild-type UBA1 or catalytically inactive UBA1 was incubated with 10 μM of Ub-Probe1-3 for 4 hours at room temperature. As we expected, Ub-Probe3-treated wtUBA1 resulted in the formation of a higher molecular weight band of the UBA1-Ub-Probe3 complex accompanied by the disappearance of UBA1 band. Catalytically inactive C632A UBA1 did not produce higher molecular weight species, indicating that the catalytic cysteine of UBA1 reacts with the dehydroalanine of Ub-Probe3. It is important to note here that the dehydroalanine in the Ub-Probe3 is located three atoms away from the carbonyl carbon in the natural Ub-AMP adduct. The probe also lacks one of the non-bridging oxygens of the natural phosphate motif. These differences might reduce the binding affinity and E1 labeling efficiency of Ub-Probes. However, this drawback can be overcome by increasing the binding affinity of the AMP moiety by modulating the N^6-substituent of AMP. In accordance with this, Ub-Probe2 that contains propargyl substituent at the N^6-position of the adenine was less potent than Ub-Probe3 in labelling UBA1, which is also in agreement with our previous study.[22] Notably, Ub-Probe1, which lacks an AMP moiety, did not induce any visible covalent labelling of UBA1, suggesting that the occupancy of both UBL and ATP binding sites in E1s is required to efficiently induce the conformational changes of E1s for the nucleophilic attack.[9, 10] Alternatively this could be because the binding affinity of Ub-Probe1 is too low to allow the covalent bond formation with the catalytic cysteine of UBA1. Taken together, our approach allows

ChemComm Accepted Manuscript

This journal is © The Royal Society of Chemistry 20xx
Please do not adjust margins

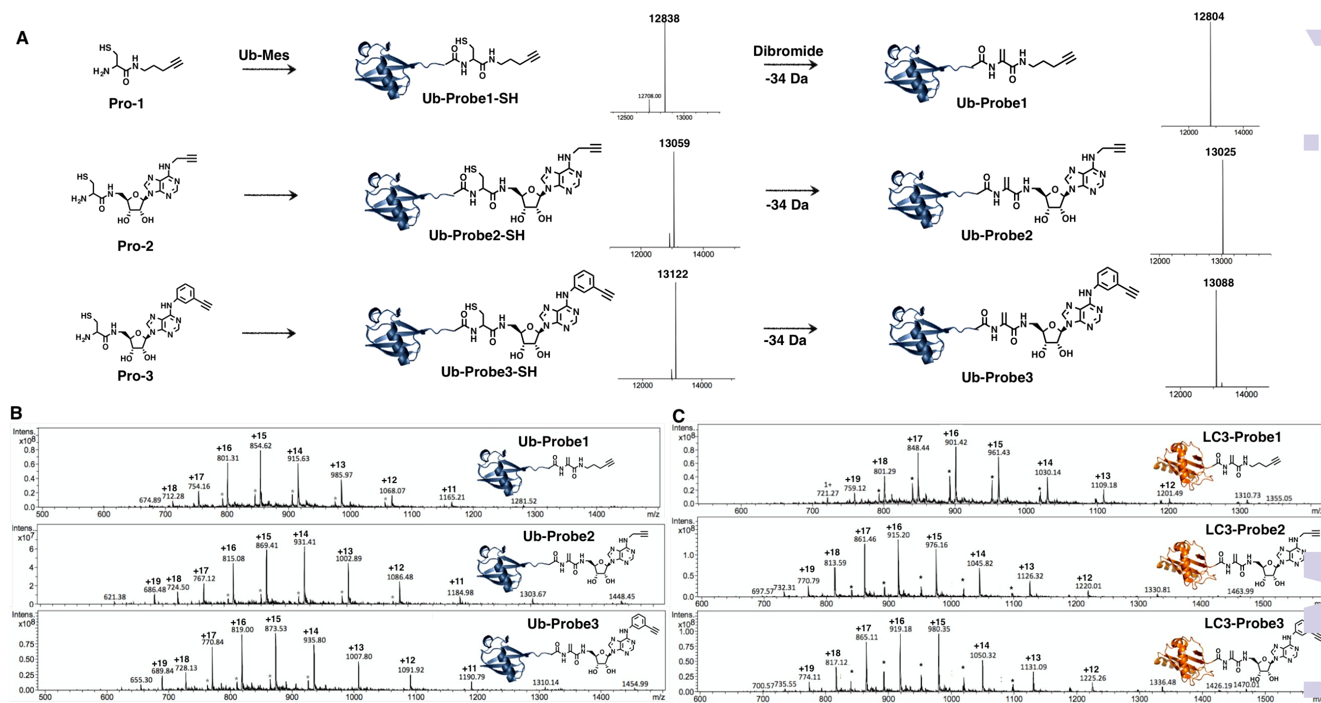


Figure 2 (A) Synthesis of Ub-Probes. Pro 1-3 were incubated with Ub~Mes for 2 hours, and the formation of Ub-Probe-SH was confirmed by LC-MS. Incubation with di-bromide reagent led to the formation of the desired probes as confirmed by LC-MS. (B) Chromatogram of the final Ub-probes with different charge states marked. Grey stars: around 10 % of the engineered protein underwent N-terminal methionine excision (NME) in *E. coli* during Ub~Mes preparation (Fig S1). This product should behave exactly the same as non-NME product. (C) Chromatogram of the final LC3-probes with different charge states marked (Fig. S2-S5). Black stars: unmodified LC3

facile synthesis and structure activity relationship study of diverse electrophilic UBL-AMP mimic probes.

Next, we investigated the reaction between ATG7 and electrophilic LC3-probes (Fig. 3B). ATG7 was incubated with LC3-Probe1-3 for 12 hours. The reaction mixture was then resolved by SDS-PAGE, followed by coomassie staining. Similar to the Ub-probes, the appearance of a higher molecular weight band was observed accompanied by the consumption of ATG7, suggesting the formation of the covalent ATG7-LC3-probe adducts. Interestingly, electrophilic LC3-Probe2 that contains propargyl group was the most efficient at labelling ATG7, indicating that the N^6-position of the adenine ring can be used to design selective electrophilic probes for various E1 enzymes. As expected, LC3-probe2 did not label UBA1, showing that LC3-probe2 is a selective mechanism-based probe for ATG7. During our studies, we also found that the dehydroalanine can react with commonly used reducing agents, such as β-mercaptoethanol and TCEP, thus caution is required when

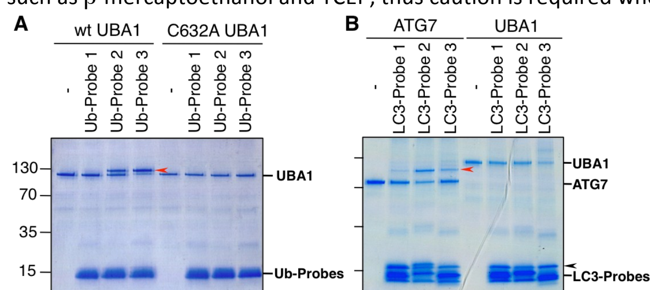


Figure 3. Labelling of E1s with UBL-Probes (A) Incubation of Ub-Probes with wtUBA1 or catalytically inactive UBA1 suggested that the catalytic cysteine is required for labeling and that Ub-Probe3 is the most potent probe. (B) LC3-Probe2 showed the strongest labeling efficiency. LC3-Probes did not show any reactivity toward UBA1.

using these reagents (Fig. S6 and S7). We found that our probes tolerated at least 100 μM TCEP.

Another advantageous feature of our probes is the presence of an alkyne tag. Conjugation of a rhodamine dye (Rh) to alkyne-tagged protein complexes via copper mediated [3+2] cycloaddition (hereafter click-chemistry) enables more direct and accurate detection and quantification of UBL-probes and their covalent adducts. Indeed, click-chemistry-mediated detection method enabled us to investigate the direct correlation between the UBA1-Ub-Probe3 complex formation and the inhibition of UBA1 (Fig. 4A). Treatment of UBA1 with Ub-Probe1 or Ub-Probe3 for 4 hours was followed by the addition of UbcH5a (E2), Rsp5 (E3), GFP-Sic60 (substrate), Ub, and ATP. Subsequent conjugation to Rh-azide and dual-colour in-gel fluorescence scanning allowed simultaneous monitoring of the ubiquitination of GFP-Sic60 and the covalent labelling of E1. The appearance of a Rh-fluorescent band (red) in the 130 kDa region indicated the covalent labelling of UBA1 with Ub-Probe3, which in turn reduced catalytic activity of UBA1 as judged by the decreased ubiquitination of GFP-substrate (green). Ub-Probe1 treated UBA1 did not show a labelling of UBA1 nor a decrease in substrate ubiquitination levels, confirming that Ub-probe1 is not a potent inhibitor of UBA1 (Fig. 4A).

Since deubiquitinating enzymes (DUB) possess strong nucleophilic cysteines that selectively target ubiquitinated substrates, we investigated the reactivity of Ub-probes towards DUBs using click-chemistry (Fig. 4B). We also tested the labelling selectivity within the E1 enzyme family using SUMO E1. Thus, Ub-Probe1 or Ub-Probe2 was incubated with UBA1, SAE1/2 (SUMO E1) or IsoT (DUB). As expected, Ub-Probe2 covalently labelled UBA1 but not SUMO E1, highlighting the selectivity of Ub-Probe2 (Fig. 4B, S8). However, incubation of IsoT with Ub-probe1-2 induced almost no

ChemComm Accepted Manuscript

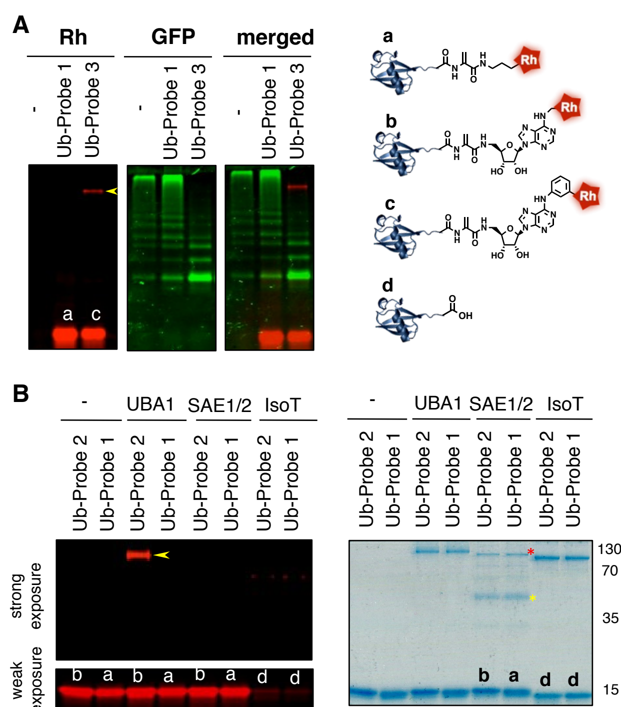


Figure 4. The alkyne tag provides an additional detection handle. (A) The labeling of UBA1 and its activity change was detected by dual-colour in-gel fluorescence scanning. Yellow arrow: UBA1-Ub-Probe3 adduct (B) Ub-Probe2 labels UBA1 but not SUMO E1. IsoT DUB cleaved the amide bond in Ub-Probes. Yellow star: SAE1. Red star: SAE1/2.

detectable labelling when compared with the labelling of UBA1 (Fig. 4B, S8). Unexpectedly, the fluorescence signals of free Ub-Probe1 and Ub-Probe2 disappeared upon IsoT treatment (Fig. 4B). We hypothesize that this is because IsoT cleaves the amide bond in Ub-Probes rather than forming a covalent bond with the neighbouring dehydroalanine (Fig. 4B). This highlights the remarkable chemical precision of IsoT active site. This discovery was only possible using the click-chemistry-based detection methods.

In summary, we reported a facile synthesis of mechanism-based electrophilic UBL-AMP mimic probes for E1 enzymes. The resulting probes can covalently and selectively react with their cognate E1s, leading to the formation of protein complexes that resemble tetrahedral intermediate of E1-UBL-AMP complex. We showed that our synthetic approach is applicable to the synthesis of UBL-probes (specifically Ub and LC3 based probes) that target both canonical and non-canonical classes of E1s. We envision that the developed synthetic approach will be generally applicable to other UBL based probes, since all UBL proteins have either no cysteine or one cysteine residue except FAT10 and SUMO, which allows dehydroalanine chemistry. Furthermore, we showed that the alkyne tag in UBL electrophilic probes provides an additional detection handle, enabling an accurate analysis and quantification of proteins labelled by the probes. This can be used to evaluate the efficiency of the covalent labelling of E1 enzyme with different probe analogues at N^6-position of the adenine ring to identify the most efficient electrophilic probe for subsequent crystallography studies. Further studies aimed at using electrophilic UBL-AMP mimic probes to study conformational rearrangements of E1 enzymes during the catalytic cycles will be reported in the near future.

We thank the Schulman lab for providing the LC3-Mes, and David Krist for proofreading the manuscript. Funding from Northwestern University is greatly acknowledged. A. V. S. is a Pew Scholar in the Biomedical Sciences, supported by the Pew Charitable Trusts.

1. M. Hochstrasser, *Nature*, 2009, **458**, 422-429.
2. F. C. Streich, Jr. and A. L. Haas, *Sub-cellular biochemistry*, 2010, **54**, 1-16.
3. B. A. Schulman and J. W. Harper, *Nature reviews. Molecular cell biology*, 2009, **10**, 319-331.
4. H. Walden, M. S. Podgorski and B. A. Schulman, *Nature*, 2003, **422**, 330-334.
5. S. K. Olsen and C. D. Lima, *Molecular cell*, 2013, **49**, 884-896.
6. L. M. Lois and C. D. Lima, *The EMBO journal*, 2005, **24**, 439-451.
7. N. N. Noda, K. Satoo, Y. Fujioka, H. Kumeta, K. Ogura, H. Nakatogawa, Y. Ohsumi and F. Inagaki, *Molecular cell*, 2011, **44**, 462-475.
8. J. P. Bacik, J. R. Walker, M. Ali, A. D. Schimmer and S. Dhe-Paganon, *The Journal of biological chemistry*, 2010, **285**, 20273-20280.
9. S. K. Olsen, A. D. Capili, X. Lu, D. S. Tan and C. D. Lima, *Nature*, 2010, **463**, 906-912.
10. X. Lu, S. K. Olsen, A. D. Capili, J. S. Cisar, C. D. Lima and D. S. Tan, *Journal of the American Chemical Society*, 2010, **132**, 1748-1749.
11. P. E. Dawson, T. W. Muir, I. Clark-Lewis and S. B. Kent, *Science (New York, N.Y.)*, 1994, **266**, 776-779.
12. J. M. Chalker, S. B. Gunnoo, O. Boutureira, S. C. Gerstberger, M. Fernandez-Gonzalez, G. J. L. Bernardes, Griffin, H. Hailu, C. J. Schofield and B. G. Davis, *Chemical Science*, 2011, **2**, 1666-1676.
13. N. Haj-Yahya, H. P. Hemantha, R. Meledin, S. Bondalapati, M. Seenaiah and A. Brik, *Organic letters*, 2014, **16**, 540-543.
14. When the same starting material is used, preparation of AVSN takes 10 synthetic steps while our approach requires 2 steps.
15. K. D. Wilkinson, T. Gan-Erdene and N. Kolli, *Methods in enzymology*, 2005, **399**, 37-51.
16. O. N. Burchak, M. Jaquinod, C. Cottin, L. Mugherli, K. Iwai, F. Chatelain and M. Y. Balakirev, *Chembiochem : a European journal of chemical biology*, 2006, **7**, 1667-1669.
17. F. El Oualid, R. Merkx, R. Ekkebus, D. S. Hameed, J. J. Smit, A. de Jong, H. Hilkmann, T. K. Sixma and H. Ovaa, *Angewandte Chemie (International ed. in English)*, 2010, **49**, 10149-10153.
18. F. Janku, D. J. McConkey, D. S. Hong and R. Kurzrock, *Nature reviews. Clinical oncology*, 2011, **8**, 528-539.
19. W. K. Wu, S. B. Coffelt, C. H. Cho, X. J. Wang, C. W. Lee, F. K. Chan, J. Yu and J. J. Sung, *Oncogene*, 2012, **31**, 939-953.
20. T. Shpilka, H. Weidberg, S. Pietrokovski and Z. Elazar, *Genome biology*, 2011, **12**, 226.
21. P. Ma, O. Schillinger, M. Schwarten, J. Lecher, R. Hartmann, M. Stoldt, J. Mohrluder, O. Olubiyi, B. Strodel, D. Willbold and O. H. Weiergraber, *Biochemistry*, 2015, **54**, 5469-5479.
22. H. An and A. V. Statsyuk, *Chemical Science*, 2015, **6**, 5235-5245.

ChemComm Accepted Manuscript

This journal is © The Royal Society of Chemistry 20xx