# Non-canonical scaffold-type ligase complex mediates protein UFMylation

Joshua J. Peter[1], Helge M. Magnussen[1], Paul Anthony DaRosa[2], David Millrine[1], Stephen P Matthews[1], Frederic Lamoliatte[1], Ramasubramaniam Sundaramoorthy[3], Ron R Kopito[2] and Yogesh Kulathu[1]*

**Affiliations –**

[1] Medical Research Council Protein Phosphorylation & Ubiquitylation Unit (MRC-PPU), School of Life Sciences, University of Dundee, Dundee DD1 5EH, UK.

[2] Department of Biology, Stanford University, Stanford, CA 94305, USA

[3] Centre for Gene Regulation and Expression, School of Life Sciences, University of Dundee, Dundee DD1 5EH, UK

***corresponding author –**

Dr Yogesh Kulathu, Medical Research Council Protein Phosphorylation & Ubiquitylation Unit (MRC-PPU), School of Life Sciences, University of Dundee, Dow Street, Dundee DD1 5EH, UK.

Tel: +44 1382 388163; E-mail: ykulathu@dundee.ac.uk

Running Title: Mechanism of UFMylation

## Abstract

Protein UFMylation is emerging as a posttranslational modification essential for endoplasmic reticulum and cellular homeostasis. Despite its biological importance, we have a poor understanding of how UFM1 is conjugated onto substrates. Here, we use a rebuilding approach to define the minimal requirements of protein UFMylation. We find that the reported E3 ligase UFL1 is inactive on its own and identify UFBP1 to bind UFL1 to form an active E3 ligase complex. While UFC1 is an intrinsically Cys-reactive E2, we do not identify any catalytic cysteines on UFL1/UFBP1, suggesting a scaffold-type E3 ligase mechanism. Interestingly, the E3 ligase complex consists of winged-helix (WH) domain repeats that activates UFC1 for aminolysis. We identify the adaptor protein CDK5RAP3 to bind to and regulate E3 ligase activity potentially by preventing off-target UFMylation. In summary, our work identifies the minimal requirements for UFMylation and reveals regulatory principles of this atypical E3 ligase complex.

UFM1 is a ubiquitin-like (Ubl) modifier that contains the characteristic β-grasp fold found in all UBLs and is highly conserved in multicellular eukaryotes [1,2]. Like other UBLs, UFM1 is covalently attached to lysine residues on substrates via an enzymatic cascade involving an E1 activating enzyme, UFM1-activating enzyme 5 (UBA5), E2 conjugating enzyme, UFC1, and an E3 ligase, UFL1 (UFM1 E3 Ligase 1) [3,4]. Recent studies have identified the ribosomal protein RPL26 to be the principal target of UFMylation which plays a critical role in endoplasmic reticulum (ER) homeostasis [5–7]. Besides ER associated roles, UFMylation has also been implicated in other cellular processes including protein translation, DNA damage response and nuclear receptor mediated transcription [8–13]. Further, loss of or mutation of components of the UFMylation machinery has been linked to many diseases such as cancer, type-2 diabetes, neurological disorders and cerebellar ataxia, where failure of ER homeostasis and protein quality control could be one of the major contributing factors [12,14–17]. Indeed, UFMylation is an essential posttranslational modification for animal development as loss of UFM1 or any of the UFMylation enzymes results in failure of erythropoiesis and embryonic lethality [18–21].

UFM1 is expressed in a precursor form consisting of 85 amino acids. It is then post-translationally processed by UFM1-specific proteases (UFSPs) which cleave off the last two residues at the C-terminus to generate mature UFM1 that contains an exposed C-terminal glycine, $Gly^{83}$ [22]. Mature UFM1 is activated by the E1, UBA5 via the formation of a high energy thioester bond between $Cys^{250}$ of UBA5 and $Gly^{83}$ of UFM1 [3,23]. Activated UFM1 is then transferred from UBA5 on to the catalytic $Cys^{116}$ of UFC1 [3,24]. UFC1 has a canonical UBC fold and in addition an N-terminal helical extension whose function is not known [25,26]. UFC1 lacks many of the features conserved in many E2s such as the catalytic HPN motif which suggests unique modes of action and regulation. Further, ubiquitin E2 enzymes such as UBE2D3 associate with ubiquitin via a backside interaction which mediates E2 dimerization and processivity of polyubiquitin formation [27,28]. For UFC1, it is unknown if similar mechanisms exist to regulate its activity as detailed biochemical characterization is lacking.

UFL1 is the only E3 ligase identified to date in the UFMylation pathway catalysing the final step of transfer of UFM1 from UFC1 to substrates [29]. Indeed, loss of UFL1 results in loss of UFMylation and mice lacking Ufl1 phenocopy loss of UFM1

exhibiting failure in haematopoiesis and embryonic lethality suggesting that it could be the main E3 ligase [30]. While data implying UFL1 is the E3 ligase stem from knockout studies and overexpression studies, direct biochemical evidence demonstrating that UFL1 is an E3 ligase is insufficient [24]. Consequently, the mechanism of how UFL1 functions as an E3 ligase is also not understood. In general, E3 ligases are classified either as scaffold-type ligases that bring together a UBL-charged E2 with substrate for direct transfer of the UBL from the E2 on to substrate, or as Cys-dependent enzymes where the UBL is first transferred from the E2 to the catalytic Cys of the E3 for subsequent transfer to the substrate [31]. Examples of scaffold-type E3s are RING family ligases and RANBP2 whereas HECT, RBR and RCR enzymes typify the latter type [32–36]. Intriguingly, UFL1 does not possess any conserved sequence or domain features found in known E3 ligases. Hence it is not clear if UFL1 is a Cys-dependent E3 enzyme or uses a scaffolding mechanism.

UFL1 is predominantly found anchored at the ER through its interaction with an adaptor protein, DDRGK1/UFBP1 (UFM1 binding protein-1) which has also been suggested to be one of the main substrates of UFL1 [29,37]. UFBP1 localizes to the ER membrane via an N-terminal transmembrane segment and its function is unclear [5]. Intriguingly, loss of UFBP1 affects the stability and expression levels of UFL1 [20,29,37]. CDK5RAP3 is another poorly characterized protein commonly associated with UFL1. Several anecdotal reports have led to speculations that UFL1, UFBP1 and CDK5RAP3 together form an E3 ligase complex [24,38]. Only a handful of substrates of UFL1 have been identified to date and many of them are ER localized [38,39]. In addition to monoUFMylation, polyUFM1 chains can be assembled which are mainly linked via K69 [12]. In the ubiquitin system, the type of polyUb linkage formed is generally dictated by the last enzyme in the cascade that forms a thioester linkage with ubiquitin before transfer to substrate. Scaffold-type E3s such as RING E3 ligases bind to charged E2 (E2~Ub) to mediate transfer of Ub onto susbtrate, and hence the E2 enzymes are thought to determine the linkage type assembled. In contrast, catalytic Cys containing E3 ligases form an E3~Ub thioester intermediate and dictate the linkage type formed [31,34].

Given the lack of understanding of the molecular mechanism underpinning the UFM1 conjugation machinery and the mechanisms governing its regulation, here we adopt

a rebuilding approach with purified components to fully reconstitute UFMylation *in vitro*. Using this reductionist approach, we reveal that UFL1 on its own is an unstable protein that cannot support UFMylation. We demonstrate that UFL1 together with UFBP1 forms a functional E3 ligase complex and provide in depth biochemical characterization of the ligase complex. Structure prediction using Alphafold provides an explanation for the underlying reason for complex formation between UFL1 and UFBP1, and offers a rational framework for making truncations which help establish the minimal regions on UFL1 and UFBP1 essential for ligase activity. Parallel investigations of UFC1 reveal that it is an intrinsically Cys reactive enzyme that is activated by the UFL1/UFBP1 ligase complex for aminolysis. Our analyses reveal UFL1/UFBP1 to be a scaffold-type E3 ligase. Further, we identify CDK5RAP3 to bind to the E3 ligase complex and inhibit UFMylation by preventing discharge of UFM1 from UFC1. Importantly, by reconstituting UFMylation of ribosomes *in vitro,* we find that CDK5RAP3 has a regulatory function as it restricts UFMylation towards *bona fide* substrates and lysine selectivity. Lastly, we identify an N-terminal helical extension on UFC1 to have an inhibitory role that is essential for CDK5RAP3 to restrain UFMylation. Our mechanistic insights describe an atypical E3 ligase complex and provide a foundation for investigating the catalytic mechanism, regulation, and substrate specificity of this E3 ligase complex.

## Results

### UFC1 has intrinsic cysteine reactivity

The interaction of UFC1 with UBA5 and how UFM1 is transferred from UBA5 to the catalytic Cys is reasonably well understood [23,24,40–42]. We therefore focussed on subsequent events and first analysed the intrinsic reactivity and stability of UFC1~UFM1 (~ denotes the thioester bond between UFM1 G83 and UFC1 C116) in comparison to the well characterized ubiquitin E2s, UBE2D3 and UBE2L3 [43]. Intrinsic reactivity of E2s assayed using free amino acids can give insights into the ability of E2s to transfer UBLs onto substrates and can provide clues about the type of E3 ligase they work with and even the nature of the substrate [27,36]. For instance, E2s such as UBE2D3 which function together with RING/scaffold like E3 ligases are capable of aminolysis and are therefore reactive towards the free amino acid Lys (**Fig 1A, B**) [43]. On the other hand, UBE2L3 which function only with Cys-dependent E3 ligases is incapable of aminolysis and lack reactivity to free Lys on its own (**Fig 1B**) [43]. When discharge of UFM1 from UFC1~UFM1 onto free amino acids was analysed, we observed preferential discharge onto Cys while no Lys discharge was observed (**Fig 1C**). Under the same experimental conditions, UBE2D3 discharged Ub in the presence of both Cys and Lys (**Fig 1C, Fig S1**). Further, using increasing concentrations of free Lys or with prolonged incubation in a time course, we find that UFC1 has negligible Lys reactivity (**Fig 1D-E**). These results suggest that UFC1 has intrinsic Cys reactivity and may work with Cys-dependent E3 ligases.

### UFL1 together with UFBP1 forms an active E3 ligase complex

To analyse how UFC1 works together with the E3 ligase UFL1 to transfer UFM1, we aimed to reconstitute UFMylation using UFL1. We expressed 6xHis-UFL1 in *E. coli* and purified it by affinity chromatography as the first step. Subsequent size exclusion chromatography (SEC) analysis indicated UFL1 to elute in the void fraction suggesting formation of soluble aggregates (**Fig S2A**). Our attempts to enhance solubility by incorporating different solubility enhancing tags at the N-terminal or C-terminal ends of UFL1 or with different buffers and additives were unsuccessful. We therefore reasoned that UFL1 might require additional binding partners for its stability and activity. Indeed, several studies have shown that expressing subunits of protein complexes individually may result in aggregation [44,45]. To identify binding partners of

UFL1, we performed Yeast Two Hybrid (Y2H) screening using a cDNA library derived from human placenta as prey (**Fig S2B**). Surprisingly, only three hits were identified with UFM1 binding protein 1 (UFBP1, a.k.a DDRGK1) being the most significant interactor. We therefore wondered if co-expressing UFBP1 with UFL1 might confer solubility and stability to UFL1. To achieve this, we used a co-expression system to express full-length UFL1 together with UFBP1 lacking its transmembrane sequence and purified the ligase complex with three enrichment steps (**Fig S2C**). Indeed, UFL1 co-expressed with UFBP1 lacking its N-terminal transmembrane segment (UFL1/UFBP1) is soluble, well behaved and no longer formed soluble aggregates in SEC (**Fig 2A**). We confirmed the formation of a stable UFL1/UFBP1 complex by analysing its precise molecular mass using a mass photometer [46] and SEC-MALS (**Fig 2B, S2D**). This revealed UFL1/UFBP1 to have a mass of 126 kDa which corresponds to that of a heterodimeric complex. Of note, the mass photometry measurements are performed at low nanomolar concentrations highlighting the stable nature of the complex.

Having obtained soluble UFL1/UFBP1 complexes, we next tested the catalytic activity of UFL1 expressed alone and in complex with UFBP1 using *in vitro* UFMylation assays containing UBA5, UFC1 and UFM1. Whereas UFL1 on its own did not show any appreciable activity, the UFL1/UFBP1 complex is an active ligase as evidenced by the formation of multiple UFMylated products and UFL1 autoUFMylation (**Fig 2C**). To check if UFL1/UFBP1 was capable of UFMylating substrates *in vitro*, we used TRIP4/ASC1, a protein reported to be UFMylated in cells [12]. Indeed, UFL1/UFBP1 can efficiently UFMylate TRIP4 *in vitro* (**Fig 2C**). Intriguingly, exogenous addition of purified UFBP1 to UFL1 expressed on its own does not change the elution profile of UFL1 to match that of co-expressed UFL1/UFBP1 or in its E3 ligase activity (**Fig S2E-F**). Taken together, these results show that UFL1 and UFBP1 form an obligate heterodimer required for E3 ligase activity.

Since polyUFMylated products are formed in UFMylation reactions, we next analysed the reaction products by mass spectrometry to determine the linkage types assembled. Interestingly, LC/MS analyses revealed that polyUFM1 chains are formed which are predominantly linked via K69 with some K3 and K7 linkages also

observed (**Fig 2D**). To validate these observations, we generated a series of K-R mutants where every Lys in UFM1 was individually mutated to Arg. While mutation of K3, K7, K19, K34 or K41 to Arg did not affect overall UFMylation, UFM1 K69R resulted in near complete loss of polyUFMylated species and unanchored diUFM1, closely resembling reactions with UFM1 K0 in which all Lys were mutated to Arg (**Fig 2E** lanes 7&8**, S2G**). Interestingly, western blotting for UFL1 revealed only monoUFMylated UFL1 in the presence of UFM1 K69R suggesting that UFL1 autoUFMylation involves formation of polyUFM1 chains on UFL1 that are linked via K69 (**Fig 2E**). We next generated $K_{only}$ UFM1 mutants that contain only a single Lys while all others are mutated to Arg. UFMylation assays using these mutants confirm that K69 is the preferred linkage type assembled (**Fig 2F**). Single Lys-retaining mutants all abrogate polyUFMylation formation comparably to UFM1 K0 with the sole exception of K69-only UFM1 (**Fig 2F**). In reactions with UFM1 K69R or K0, mainly monoUFMylated products corresponding to modified UFL1, UFBP1 and UFC1 are formed. In summary, our reconstitution experiments reveal that UFL1 together with UFBP1 forms an active E3 ligase complex that can efficiently UFMylate substrates and assemble polyUFM1 chains that are K69-linked.

**UFL1/UFBP1 is a scaffold type E3 ligase complex activating UFC1 for aminolysis**

Our analyses find that UFC1 is an intrinsically Cys-reactive E2 suggesting that UFL1 may be a Cys-dependent enzyme (**Fig 1**). The N-terminal region of UFL1 contains four Cys and sequence analysis show that all four Cys vary in their degree of conservation with C32 being the most conserved (**Fig 3A**). Based on analysis of predicted folding propensity (**Fig S3A**) and secondary structure prediction, we made a C-terminal truncation of UFL1, 1-410, which when expressed on its own formed soluble aggregates like full length UFL1 and required co-expression of UFBP1 to obtain a heterodimeric complex (**Fig S3B**). Importantly, this $UFL1^{1-410}$/UFBP1 complex is an active ligase **(Fig S3C**). To identify the catalytic Cys in UFL1, we mutated each Cys individually to Ala in the $UFL1^{1-410}$/UFBP1 complex. To our surprise, *in vitro* UFMylation assays showed that none of the single mutants had any impact on the autoUFMylation activity of UFL1 or UFMylation of substrates implying that UFL1 does not contain a catalytic Cys (**Fig 3B**). Since UFBP1 does not have any Cys residues, these results suggest that the UFL1/UFBP1 ligase complex may

instead employ a scaffolding mechanism to transfer UFM1 on to substrate. Further, we do not observe any transthiolation products of UFL1/UFBP1 **(Fig S3D)**.

This raises the possibility that the UFL1/UFBP1 ligase complex could induce aminolysis of UFC1~UFM1. Hence, we compared discharge of UFM1 from UFC1~UFM1 onto Lys in the absence and presence of the ligase complex. Whereas UFC1 does not discharge UFM1 onto Lys on its own, UFM1 was readily discharged onto Lys in the presence of UFL1/UFBP1 (**Fig 3D**). Based on these observations, we suggest that UFL1/UFBP1 functions as a scaffold-type E3 ligase that binds to charged UFC1 to promote aminolysis.

While inefficient, UFC1 on its own can assemble free UFM1 chains which is significantly enhanced in the presence of the UFL1/UFBP1 ligase complex (**Fig S3E-F**). Interestingly, UFC1 assembles mainly K69 linkages (**Fig 3D**) and this linkage specificity is maintained in the presence of UFL1/UFBP1 (**Fig 2D**). Of note, this reveals that the linkage specificity is determined by the E2 and is not altered by the E3 ligase complex, a feature commonly observed in ubiquitin RING E3 ligases [34,47,48]. These results further strengthen our conclusion that UFL1/UFBP1 is a scaffold-type ligase complex.

**Tandem WH domains of UFL1/UFBP1 constitute minimal E3 ligase**

As UFL1 does not have any obvious sequence or domain features found in any of the known E3 ligases, we attempted to define the minimal catalytic region on UFL1. As our efforts to determine the structure of this complex were not successful, we used the recently released Alphafold pipeline to predict the structure (**Fig 4A**) [49]. Structure prediction of the complex using Multimer had good PAE scores (**Fig S4A**) and shows UFL1 and UFBP1 form a heterodimer (**Fig 4A-B**). Interestingly, UFL1 is predicted to contain 5 winged-helix (WH) domain repeats and a C-terminal helical domain (**Fig 3A** blue, S4B). The WH repeats are arranged such that a helix from each WH domain forms a helical backbone. The fifth WH domain contains a stack of $\alpha$-helices at its C-terminus. These WH domains could be classified as PCI-like WH domains, named so because it is commonly found in components of the proteasome lid, the COP9 signalosome and the eukaryotic translation initiator, eIF3 [50,51]. In addition, UFL1 is predicted to have an N-terminal helix (1-25) followed by a partial or hemi-WH domain (pWH) (**Fig 4B,** blue). Following the transmembrane segment at its

N-terminus, UFBP1 has a long coiled-coil region (NTR) followed by a WH (WH') domain and ends with a hemi-WH domain (pWH') (**Fig 4B, teal**). Interestingly, the hemi-WH (pWH) from UFL1 complements a hemi WH (pWH') from the C-terminus of UFBP1 to form a complete WH domain (**Fig 4B**), thus providing an explanation to why UFBP1 is required to stabilize UFL1. Superposition of the WH domains of UFL1 and UFBP1 reveal them to have identical folds except for a β-strand which is lacking in the composite WH domain (**Fig S4C-D**). The coiled-coil nature of UFBP1 and the helical backbone of UFL1 may explain the unusual migration of UFL1/UFBP1 by SEC. To validate the predicted structure of the UFL1/UFBP1 complex, we deleted the N-terminal helix along with pWH of UFL1 and checked for complex formation with UFBP1. No soluble protein was obtained upon deletion of this region of UFL1. Likewise, removal of C-terminal pWH domain of UFBP1 also resulted in poor expression, highlighting the importance of the formation of the composite WH (pWH-pWH') domain for complex formation and protein stability.

Guided by the insights from the Alphafold predictions, we made additional truncations to map the minimal catalytic domain of the ligase complex (**Fig 4C**). The truncated complexes expressed well and were purified to homogeneity (**Fig S3F**). We first analysed the ability of the different truncated complexes to promote discharge of UFM1 from UFC1~UFM1 onto Lys. This revealed that all the truncations tested could discharge UFM1 and the smallest region able to promote discharge contains one WH domain each from UFL1 and UFBP1 and the composite WH domain (**Fig 4D,S5A-B**).

We then monitored the ability of the different UFL1/UFBP1 complexes to bind to UFC1 for which we performed analytical SEC of the different UFL1/UFBP1 constructs with UFC1 on its own or UFC1$^{\text{-O-UFM1}}$ where UFM1 is linked to C116S via an oxyester bond (**Fig S5 C-D**). UFC1$^{\text{-O-UFM1}}$ is more stable compared to thioster-linked UFC1~UFM1 and therefore used for these analyses. Post preparation, we incubated UFC1$^{\text{-O-UFM1}}$ with UFSP2 or 0.2N NaOH and observed complete collapse only in alkaline conditions confirming that the purified UFC1$^{\text{-O-UFM1}}$ is indeed linked via an oxyester linkage **(Fig S4E)**. Analysis of complex formation revealed that UFL1/UFBP1 does not bind to UFC1 or UFM1 on its own but forms a stable complex with UFC1$^{\text{-O-UFM1}}$ that is of high enough affinity to elute as a complex on SEC (**Fig 4E**). Interestingly, even the smallest of the UFL1/UFBP1 variants was able to bind to

UFC1$^{-O-UFM1}$ (**Fig 4F**). Taken together, the ligase complex containing the N-terminal segment of UFL1 and the C-terminal region of UFBP1, which we refer to as UFL1/UFBP1$^{min}$ (denoted as complex I in **Fig 4C**) is sufficient for binding to charged UFC1 and may represent the minimal ligase domain.

Next, we assayed ligase activity using UFMylation assays monitoring diUFM1 formation. Like full length UFL1, the different C-terminal truncations co-expressed with UFBP1 did not exhibit any loss of activity. Thus, ligase activity requires the N-terminal region of UFL1. Similarly, truncation of the N-terminus region (NTR) of UFBP1 did not affect diUFM1 formation. This suggests that the minimal ligase domain is sufficient to form diUFM1 chains. Lastly, we checked the impact of truncations of different regions on UFL1/UFBP1 complex on their ability to modify the substrates such as Histone H4 and MRE11 (**Fig 4G, S4G**). Intriguingly, while deletion of regions outside of WH1, WH2 did not alter diUFM1 formation, these deletions affected substrate modification. Further, deletion of the N-terminal region (NTR) of UFBP1 significantly impacted substrate UFMylation suggesting roles for these regions of the ligase complex in substrate recognition **(Fig 4G-H)**. Hence, using systematic truncations based on sequence and Alphafold predictions, we here make the surprising discovery that three tandem WH domains of UFL1/UFBP1 mediate E3 ligase activity.

**CDK5RAP3 restricts the E3 ligase activity of UFL1/UFBP1**

Our Y2H screen identified a second hit with low significance, CDK5 regulatory subunit-associated protein 3 (CDK5RAP3, also known as LZAP), an evolutionarily conserved 53 kDa protein that lacks any known functional domains or motifs. Secondary structure analysis predicts CDK5RAP3 to contain a central unstructured region flanked by $\alpha$-helical stretches at the N and C terminus (**Fig S5**). To determine if CDK5RAP3 can interact with UFL1 in the context of the UFL1/UFBP1 complex, we incubated recombinant full length CDK5RAP3 with UFL1/UFBP1 and analysed complex formation by SEC. Analysis of UV chromatograms and the fractions obtained, confirmed that CDK5RAP3 interacts with UFL1/UFBP1 *in vitro* and forms a stable complex (**Fig 5A**). We further analysed the complex by mass photometry which confirmed the presence of a stable ternary complex with an experimental molecular mass of 192 kDa (**Fig 5B**).

Since CDK5RAP3 forms an integral complex with UFL1/UFBP1 and CDK5RAP3 has recently been suggested to function as a substrate adaptor [52], we wondered if it could influence E3 ligase activity or substrate UFMylation . We therefore performed *in vitro* UFMylation assays and monitored the E3 ligase activity of UFL1/UFBP1 in the presence of increasing concentrations of CDK5RAP3. Surprisingly, incubation of CDK5RAP3 with UFL1/UFBP1 impaired E3 ligase activity and UFMylation in a concentration dependent manner **(Fig 5C)**. Moreover, preassembled UFL1/UFBP1/CDK5RAP3 complexes purified by gel filtration also showed no ligase activity (**Fig S6A**). The presence of CDK5RAP3 also inhibits UFL1 autoUFMylation and MRE11 and histone H4 UFMylation (**Fig 5C,** bottom, **Fig S6E**). These observations imply that CDK5RAP3 binds to inhibit the ligase complex or restricts its activity towards specific substrates.

To gain insights into how CDK5RAP3 inhibits UFMylation, we tested if CDK5RAP3 inhibits complex formation between UFL1/UFBP1 and UFC1$^{\text{-O-UFM1}}$. When analysed by pull down and analytical SEC, we find that CDK5RAP3 forms a complex together with UFL1/UFBP1 and UFC1$^{\text{-O-UFM1}}$ (**Fig S6B-C**). Since CDK5RAP3 does not affect UFC1~UFM1 binding, we next tested if CDK5RAP3 binding influences discharge of UFM1 from UFC1~UFM1 by monitoring transfer of UFM1 onto free Lys. While UFM1 is readily discharged onto Lys in the presence of UFL1/UFBP1, this is completely blocked in the presence of CDK5RAP3 (**Fig 5D-E**). These observations suggest that CDK5RAP3 may bind to UFC1~UFM1 to inhibit the ligase complex.

While these *in vitro* experiments clearly demonstrate that CDK5RAP3 blocks UFC1 activation and UFMylation, it raises the question of the role of CDK5RAP3 in substrate UFMylation, especially UFMylation of ribosomes, the best described UFMylation substrate to date [5,6]. To address this question, we set up a cell-free reconstitution of ribosome UFMylation. Strikingly, addition of UFL1/UFBP1 to purified 60S ribosomes leads to robust modification of RPL26 (**Fig 5F**), and in line with previous observations, both mono- and di-UFMylated RPL26 are observed. When increasing concentrations of CDK5RAP3 is titrated into this UFMylation reaction, the di-UFMylation of RPL26 is completely abolished (**Fig 5F**). Importantly, even at the highest concentration of CDK5RAP3, monoUFMylation of RPL26 is not affected. Hence, we propose CDK5RAP3 to be a specificity determinant directing activity of

the ligase complex towards substrate Lys by limiting spurious and off-target UFMylation.

**Regulation of UFC1 activity by its N-terminus**

In addition to its UBC fold, UFC1 has at its N-terminus a highly conserved α-helix (α0) whose function is unknown (**Fig 6A**). Further, UFC1 lacks the oxyanion hole stabilizing Asn as part of the highly conserved HPN motif found in E2 enzymes, which is instead replaced by a TAK motif at this position (**Fig 6B**). To determine if α0 has a role in regulating UFMylation, we first compared the activity of $UFC1^{WT}$ and UFC1 lacking α0 ($UFC1^{\Delta N}$) in UFMylation assays containing UFC1 with or without UFL1/UFBP1. Surprisingly, $UFC1^{\Delta N}$ showed stronger overall UFMylation compared to $UFC1^{WT}$ in the presence of UFL1/UFBP1 (**Fig 6C**). These results suggest an inhibitory role for the N-terminal helical extension of UFC1. Next, we compared discharge of UFM1 from UFC1~UFM1 and $UFC1^{\Delta N}$~UFM1 onto free Lys. In the absence of E3 ligase, $UFC1^{\Delta N}$ rapidly discharges UFM1 onto Lys (**Fig 6D & E**). In the presence of UFL1/UFBP1, discharge of UFM1 onto Lys is enhanced when α0 of UFC1 is deleted (**Fig 6D & E)**. Taken together with the increase in UFMylation in the presence of UFL1/UFBP1, these results further cement an inhibitory role for the N-terminal helix of UFC1.

Since we identified the N-terminal α0 of UFC1 to restrain UFMylation, we wondered if inhibition by CDK5RAP3 was mediated via this helix. Indeed, UFMylation assays comparing activity of $UFC1^{WT}$ and $UFC1^{\Delta N}$ reveal that the inhibition of E3 ligase activity by CDK5RAP3 requires the N-terminal helix of UFC1 (**Fig 6D & F, S8**). Based on these results we propose that CDK5RAP3 may interact with both UFL1/UFBP1 and α0 of UFC1 to clamp the complex in an inhibited state. Together these analyses reveal a previously unappreciated regulatory role for the N-terminal helix of UFC1 in modulating UFMylation.

## Discussion

In this study, we establish a robust *in vitro* reconstitution system using purified active components of the UFM1 enzymatic pathway to reveal the minimal requirements for UFMylation and mechanistic insights into the ligase machinery. Previous reports have provided conflicting views on the roles of UFBP1 and CDK5RAP3, with some suggesting that they are mainly substrates of UFMylation [13,29]. Further, it has been suggested that UFBP1 must be UFMylated first at K267 before it can associate with UFL1 to subsequently support UFL1 ligase activity [12]. Since both UFL1 and UFBP1 complexes are purified from bacteria, our work unequivocally demonstrates that UFL1 and UFBP1 can associate in the absence of any PTMs and the formation of this complex is essential for it to function as an E3 ligase.

It has long been debated as to whether UFL1 is a scaffold/adaptor-type E3 ligase or a Cys-based HECT-like enzyme. Scaffolding type E3 ligases such as RING E3 ligases bind to both E2 and substrate to bring them together for substrate transfer. Since UFL1 can bind to both UFC1 and its supposed substrate UFBP1 at the same time, it was suggested that UFL1 maybe a scaffold-type E3 ligase [3]. More recently, a UFMylation assay that relied on UFL1 present in mammalian cell extracts to which *in vitro* generated biotinylated E2~UFM1 thioesters was added, found that UFMylation occurred even when cell extracts were treated with cysteine-alkylating reagents, further suggesting that UFL1 is a scaffold-type E3 ligase [13]. Our mutational analysis showing that the UFL1/UFBP1 complex lacks a conserved catalytic Cys combined with the ability of the E3 ligase to promote UFC1~UFM1 aminolysis establish UFL1/UFBP1 is a scaffold-type E3 ligase. Further reinforcing this conclusion is the observation that K69-linkage specificity is imparted by the E2 in ligase free di-UFM1 formation assays and this K69-linkage specificity is maintained by the E3 ligase. Indeed, this is a feature observed in RING E3 ligases where linkage-specificity is determined by the E2 enzyme [31,34]. Since UFL1/UFBP1 promotes aminolysis, we speculate that UFL1/UFBP1 binding induces a closed UFC1~UFM1 conformation akin to RING and atypical SUMO E3 ligases [53–57].

UFC1 differs from prototypical E2s since it lacks the catalytic HPN motif, has an exposed catalytic Cys with a lower pKa, lacks the C-terminal $\alpha$-helix observed in canonical E2s and has an additional $\alpha$-helix at its N-terminus ($\alpha$0) [25,40,58]. Further, it is unclear if the canonical “backside” interaction can occur in UFC1 [28,59]. These

unique features of UFC1 and the unconventional features of the ligase complex makes it likely that the UFM1 machinery employs a unique mechanism to transfer UFM1 from UFC1 onto substrate. Interestingly, the UBE2E enzymes also have an intrinsically disordered N-terminal extension that have an inhibitory role restricting Ub transfer thereby limiting polyUb chain formation [60]. Drawing analogy to RING E3 ligases whose binding induces a closed E2~Ub conformation, it remains to be seen whether E3 binding induces a closed UFM1~UFM1 conformation and whether $\alpha$0 impedes this process. Nevertheless, our data seem to suggest that $\alpha$0 has a regulatory role potentially preventing UFM1 discharge in the absence of the E3 ligase.

The recent advances in structure prediction enabled us to narrow the regions of UFL1/UFBP1 essential for ligase activity to a tandem repeat of three WH domains - a WH domain each from UFL1 and UFBP1 and the composite WH domain formed by the two proteins [49]. As this region can bind to UFC1~UFM1 and activate the E2 for UFM1 discharge, we define this to be the minimal catalytic domain. Further structural studies will be required to reveal why three WH domains are required to form a functional ligase complex and the identification of linchpin residues in UFL1/UFBP1 that play similar roles to RING E3 ligases to activate the E2 for aminolysis [54,55]. Intriguingly, the anaphase promoting complex (APC) subunit APC2 contains a WHB domain that binds the “backside” of UCBH10/UBE2C, located at a face opposite from the catalytic site and is an allosteric site in several E2s [61,62]. This backside binding of the WHB domain is specific towards UBE2C and is important for transfer of Ub onto specific substrates [62]. Our observations also raise the question of the function of the other WH/PCI domains in UFL1/UFBP1 to ligase function. One possibility is that the WH domains may mediate substrate recognition, binding to UFM1 or serve to either regulate recruitment or processivity of chain formation. Given that the WH domain of Cockayne syndrome group B (CSB) can bind to ubiquitin, it is tempting to speculate that one of the WH domains of UFL1/UFBP1 can bind to UFM1 [63]. WH domains commonly recognize nucleic acids and are typically found in transcription factors or as protein interaction domains [64]. To our knowledge, this is the first demonstration of catalytic activity mediated by WH domains [65], leading us to propose moonlighting functions for other WH domain containing proteins. It also raises the possibility that other WH domains could have E3 ligase

activity for UFMylation or other UBLs. Since UFL1 has been proposed to have nuclear functions in telomere maintenance and DNA damage response [10,11,66], one possibility is that the WH domains may mediate DNA binding. A further possibility is the use of the WH domains of UFL1/UFBP1 to recognize rRNA or translating mRNA during ribosome UFMylation.

Our *in vitro* studies show that formation of UFM1 chains and autoUFMylation is inhibited by CDK5RAP3 (**Fig 2C**). Indeed, previous studies have observed an altered pattern of UFMylation in the absence of CDK5RAP3 leading to the suggestion that it may be a substrate adaptor [52]. Interestingly, CDK5RAP3 was suggested to function as a sensor for ER stress to induce autophagic degradation of aberrant proteins formed as a result of ribosomal stalling [67]. Several multidomain and multiprotein E3 ligases such as PARKIN and Cullin Ring Ligases (CRLs) are inhibited and their activation is a carefully orchestrated multistep process [35,68]. Based on our observations, we propose that ligase complexes containing UFL1, UFBP1 and CDK5RAP3 represent an autoinhibited state. The surprising finding that ribosome UFMylation is not abolished in the presence of CDK5RAP3 but rather restricted to monoUFMylation leads us to suggest that CDK5RAP3 regulates ligase activity in the following manner: (i) in the absence of substrate, binding of CDK5RAP3 inhibits E3 ligase activity; (ii) this autoinhibition is relieved when substrates such as the 60S ribosome are encountered. This release from inhibition may involve conformational changes induced upon recognition of structural features on the ribosome by UFL1, UFBP1 or CDK5RAP3 to mediate substrate UFMylation. Such multi-layered regulation possibly prevents spurious UFMylation and ensures ribosome UFMylation only in the right context. While previous studies have suggested roles for UFMylation for UFL1-UFBP1 interaction, and UFMylation and phosphorylation to enhance UFMylation [10,12,19,29]. Our minimal reconstitution clearly demonstrates that the ligase complex assembles in the absence of any PTM. While we cannot rule out a role for phosphorylation in enhancing ligase activity, the rapid UFMylation of ribosomes by the reconstituted ligase suggests that it may be unlikely.

Our rebuilding approach provides insights into the assembly, minimal requirements, and mechanism of the UFL1/UFBP1/CDK5RAP3 ligase and reveals principles of protein UFMylation. The *in vitro* reconstitution system we describe here using purified components lays the foundation for future biochemical and structural studies

to understand the molecular mechanism of this unusual E3 ligase complex and UFMylation. Further, the cell-free UFMylation system can be applied to understand the logic of ribosome UFMylation and its relationship to ribosome quality control pathways.

**Acknowledgements**

We thank members of the Kulathu lab for critical reading of the manuscript and helpful comments. We thank the Zeqiraj lab (Univ of Leeds) and Prof Hofmann (Univ of Köln) for helpful discussions and advice. We acknowledge MRC Reagents and Services for providing valuable reagents, especially Simone Weidlich and Rachel Toth for cloning the constructs used in this study. We thank Mr. Joby Varghese for help with Mass spectrometry data acquisition and preliminary analysis. We thank Christopher Lapointe (Puglisi lab, Stanford) for valuable reagents and discussion. We also thank Juraj Ahel (IMP Vienna) for his help with generating composite foldindex profile of UFL1. This research was supported by ERC Starting grant (RELYUBL, 677623), the Lister Institute of Preventive Medicine, BBSRC (BB/T008172/1) and MRC grant MC_UU_00018/3 to YK; Grant R01GM074874 from the National Institute of General Medical Sciences and a generous gift from the Becker Family Foundation to R.R.K; NIH Epilepsy Training Program T32 NS007280 to PAD.

**Author contributions**

YK and JJP conceived the study. JJP designed, performed and analysed all the experiments. HM designed, expressed and purified UFL1/UFBP1 constructs based on Alphafold. PADR performed *in vitro* ribosome UFMylation assays. DM and SPM contributed reagents and RS performed SEC MALS analysis. YK analysed data, supervised the study and secured funding. JJP and YK wrote the manuscript with input from all the authors.

**Declaration of interests**

The authors declare no competing interests.

**Supplemental information**

Supplemental information includes Figures S1-S8 and Table S1.

## Materials and Methods

### Plasmids, Cloning and mutagenesis

The details of cDNA constructs used in this study is given in Table S1. Cloning of most of the constructs was performed by DSTT (University of Dundee). Briefly, mutagenesis was carried out using Q5 site-directed mutagenesis (NEB) with KOD polymerase (Novagen) according to manufacturer's protocol. Following mutagenesis, cDNA constructs were amplified using *E. coli* DH5α and purified using Qiagen mini prep kit (Qiagen). All cDNA constructs were verified by DNA sequencing and services, University of Dundee using DYEnamic ET terminator chemistry (Amersham Biosciences) on Applied Biosystems automated DNA sequencers.

### Recombinant protein expression and purification

Recombinant $His_6$-3C-UBA5, $His_6$-3C-UFM1$^{WT}$, $His_6$-3C-UFM1$^{Konly}$, $His_6$-3C-UFM1$^{KtoR}$, $His_6$-3C-Cys-UFM1, $His_6$-3C-UFC1$^{WT}$, His6-3C-UFC1$^{\Delta N}$ and $His_6$-3C-MRE11 were expressed in *E. coli* BL21 and purified using $Ni^{2+}$-NTA affinity chromatography as the first step. Briefly, *E. coli* BL21 cultures expressing $His_6$-tagged proteins were grown in 2xTY medium at 37ºC until $OD_{600}$ reached 0.6-0.8. Final concentration of 0.3 mM IPTG was added and the cultures were incubated at 18 °C for 16h. Cells were harvested, resuspended, and lysed in appropriate lysis buffer containing 25 mM Tris pH 8, 300 mM NaCl, 10% Glycerol, 2 mM DTT, 1 mM Benzamidine, 1 mM AEBSF, 1x protease inhibitor cocktail (Roche) by ultrasonication. Lysed cells were then clarified by centrifugation at 30000g for 30 mins at 4ºC. The clarified lysate was then incubated with pre-equilibrated $Ni^{2+}$-NTA Agarose beads (GE Life Sciences) for 2h in appropriate binding buffer containing 25 mM Tris pH 8, 300 mM NaCl, 10% Glycerol,10 mM Imidazole. The beads were then washed extensively using wash buffer containing 25 mM Tris pH 8, 300 mM NaCl, 10% Glycerol, 2 mM DTT and 20 mM Imidazole. Finally, the bound protein was eluted using elution buffer containing 25 mM Tris pH 8, 300 mM NaCl, 10% Glycerol, 2 mM DTT and 300 mM Imidazole. Wherever necessary, $His_6$-tags were cleaved off by incubating tagged proteins with PreScission protease at 4ºC overnight. A final size exclusion chromatography step was performed using Superdex 75 HiLoad™ 16/60 Superdex™ 75pg and HiLoad™ 16/60 Superdex™ 200 pg columns (GE Healthcare Life Sciences) with the mobile phase buffer as 25 mM Tris pH 8.0, 150 mM NaCl, 10%

Glycerol, 2 mM DTT. The purified proteins were then concentrated using Amicon™ Ultra 15 concentrators (MERCK Millipore) and stored in -80°C.

**Purification of GST-tagged proteins**

GST-TEV-TRIP4 was expressed in *E. coli* BL21(DE3) strain as described above. Cells were harvested and lysed in lysis buffer containing 25 mM Tris pH 7.5, 300 mM NaCl, 10% Glycerol and 2 mM DTT using ultrasonication. Pre-equilibrated Glutathione 4B-sepharose (DSTT, University of Dundee) beads were incubated with clarified lysate for 2 h. The beads were then washed with high salt buffer containing 25 mM Tris pH 7.5, 500 mM NaCl, 10% glycerol and 2 mM DTT. Further the beads were washed with low salt buffer containing 25 mM Tris pH 8, 150 mM NaCl, 10% glycerol and 2 mM DTT. The protein was then cleaved off the tag by incubation with Precision protease at 4ºC overnight. A second ion exchange chromatography step was carried out using Resource Q (6ml) (GE Healthcare Life Sciences) column with low salt buffer as 25mM Tris pH pH7.5, 150 mM NaCl, 10% glycerol,2 mM DTT and high salt buffer as 25mM Tris pH 7.5, 500mM NaCl, 10% glycerol and 2 mM DTT. A final size exclusion chromatography step was performed using HiLoad™ 16/60 Superdex™ 75pg (GE Healthcare Life Sciences) and the protein was buffer exchanged in buffer 25 mM Tris pH 7.5, 150 mM NaCl, 10% glycerol and 2 mM DTT. The purified proteins were then concentrated and stored in -80ºC.

**Co-expression and purification of UFL1-UFBP1 complex**

$His_6$-TEV-UFL1 and StrepII-C3-UFBP1 were cloned in a pET Duet1 construct and expressed in *E. coli* BL21 codon plus (RIPL) (Agilent) cells. Bacterial cultures were grown in 2xTY medium at 37°C until $OD_{600}$ reached 0.6. Final concentration of 0.3 mM IPTG was added to induce the expression and the cultures were incubated at 18°C for 16 h. Cells were harvested, resuspended in buffer containing 25 mM Tris pH 8.0, 300 mM NaCl, 2 mM DTT, 1mM Benzamidine, 1mM AEBSF, protease inhibitor cocktails (Roche). Cells were lysed by High pressure homogenization using French press (Avestin Emulsiflux C3, Wolflabs). The lysate was then clarified by Ultracentrifugation at 30000g for 30 min at 4°C. Affinity purification was carried out in two steps. In the first step, the clarified lysate (25 mM Tris pH 8.0, 300 mM NaCl, 20 mM Imidazole, 2 mM DTT) was passed through HisTrap™ FF (GE Healthcare Life Sciences) column. After binding, the column was washed with 10 column volumes (CV) of binding buffer containing 25 mM Tris pH 8.0, 300 mM NaCl, 20 mM

Imidazole, 2 mM DTT. After washing, bound proteins were eluted using buffer containing 25 mM Tris pH 8.0, 300 mM NaCl, 300 mM Imidazole, 2 mM DTT by applying a concentration gradient of Imidazole. The eluted proteins were then passed through StrepTrap™ (GE Healthcare Life Sciences) column pre-equilibrated with buffer containing 25 mM Tris pH 8.0, 300 mM NaCl, 2 mM DTT. After 2 CV of washing with binding buffer, the proteins were eluted using 25 mM Tris pH 8.0, 300 mM NaCl, 2 mM DTT and 2.5 mM Desthiobiotin. Finally, purified proteins were passed through HiLoad™ 16/60 Superdex™ 200pg (GE Healthcare Life Sciences). The purified protein was then stored in -80°C in buffer containing 25 mM Tris pH 8.0, 300 mM NaCl, 5% Glycerol, 2 mM DTT until further use.

**Preparation and labelling of IRDye800CW-UFM1**

Cys-UFM1$^{1-83}$ which contains a Cys residue upstream of M1 was purified and exchanged into a fresh buffer containing 50 mM HEPES, pH 7.5, 150 mM NaCl and 0.5 mM TCEP using CentriPure P10 columns (EMP Biotech). Protein was then mixed with IRDye® 800CW Maleimide (LI-COR®) at 5:1 molar ratio and incubated for 2 h at RT. Unreacted dye was removed by three-step buffer exchange process using CentriPure P10 columns (EMP Biotech) (once) and PD-10 Sephadex G-25 column (twice) in a sequential manner according to manufacturer's instruction. Finally, to remove residual unreacted dye material, the labelled protein was dialyzed into buffer containing 25 mM Tris pH 8, 150 mM NaCl, 2 mM DTT using dialysis membrane (Thermo Scientific) (MWCO 3000) overnight at 4°C and stored at -80°C until further use.

**Biochemical assays**

**Lysine discharge assays**

To analyse the intrinsic reactivity of UFC1, a substrate-independent single turnover discharge assay was employed. Firstly, UFC1 was charged with UFM1$^{WT}$ or labelled UFM1 by incubating 0.5 µM UBA5, 10 µM UFC1 and 20 µM UFM1 in a buffer containing 50 mM HEPES pH 7.5, 50 mM NaCl, 0.5 mM DTT and10 mM $MgCl_2$. The charging reaction was initiated by addition of 10 mM ATP and incubation of the reaction mix at 37°C for 20-30 min. To stop further charging of UFM1 onto UFC1, the reaction was quenched by addition of 50 mM EDTA and subsequent incubation at RT for 10 mins. To check for discharge, the quenched mixture was mixed with 50 mM lysine (pH 8.0) or other amino acids namely Serine, Threonine, Arginine,

Cysteine and incubated at 37°C. The reaction was stopped at each of the time points by mixing the reaction mix with SDS-PAGE loading dye without any reducing agent and run on a 4-12% SDS-PAGE under non-reducing conditions followed by Coomassie staining or visualization using LI-COR®.

In reactions involving the UBCH5C and UBCH7, 0.5 µM UBE1 was incubated with 10 µM UBCH5C/UBCH7 and 20 µM Ub in reaction buffer containing 50 mM HEPES pH 7.5, 50 mM NaCl,10 mM ATP and 10 mM $MgCl_2$ at 37°C for 20 mins. The reaction was quenched, and discharge was analysed as described above.

In discharge assays involving UFL1/UFBP1 and CDK5RAP3, the quenched reaction mix was added to a cocktail containing 1 µM UFL1/UFBP1 full length or truncated constructs and/or 1 µM or 2 µM of CDK5RAP3. Discharge was initiated by addition of 50 mM lysine (pH 8.0) in buffer containing 50 mM HEPES pH 7.5, 50 mM NaCl and 0.5 mM DTT and incubated at 37°C for indicated time duration. The reaction was stopped and analysed as described above.

**_In vitro_ UFMylation assays**

To check for UFMylation *in vitro,* 0.25 µM UBA5, 5 µM UFC1, 1 µM UFL1 and 20-30 µM UFM1 were incubated in reaction buffer containing 50 mM HEPES 7.5, 10 mM $MgCl_2$ and 5 mM ATP for 1h at 37ºC. The reaction was stopped by the addition of 3x SDS loading buffer containing a reducing agent. The reaction products were then separated on a 4-12% SDS-PAGE gel under reducing conditions and analysed by immunoblotting using indicated antibodies. In UFMylation assays involving substrates, around 1 µM-2 µM of recombinant substrates were incubated with 0.25 µM UBA5, 5 µM UFC1, 1µM UFL1 and 30 µM of UFM1 in buffer containing 50 mM HEPES 7.5,10 mM $MgCl_2$ and 5 mM ATP for 1h at 37°C. Following incubation, the reaction was stopped and analyzed using SDS-PAGE or Immunoblotting as described above.

**Pull down assays**

To analyse the interaction of the E3 ligase with the E2, pulldown assays were performed. Approximately 10 µM of untagged UFC1 or $UFC1^{-O-UFM1}$ was incubated with 5 µM of full length UFL1/UFBP1 in binding buffer containing 25 mM Tris pH 8.0, 300 mM NaCl, 2 mM DTT for 1 h at 4°C. Following incubation, 30 µl of pre-

equilibrated Streptavidin Sepharose beads (IBA Life Sciences) (50% slurry) was added and further incubated for about 1 h at 4°C. Following binding, centrifuge tubes containing the reaction mix was spun down at 500g for 3 mins and the supernatant was discarded to remove unbound proteins. To further remove weakly bound proteins, the beads were washed thrice with 1 ml of ice-cold binding buffer. Finally, the bound proteins were eluted by addition of binding buffer containing 2.5 mM Desthiobiotin (pH 8.0) and incubation at 4°C for 30 mins. The eluates were then analysed on 4-12% SDS PAGE under reducing conditions followed by coomassie staining.

**Analytical gel filtration chromatography analysis**

For analysing the interaction between UFL1/UFBP1 and CDK5RAP3, around 30 µgs of UFL1/UFBP1 and 15 µgs of CDK5RAP3 were mixed and incubated for 1 h on ice. Then, around 50 µls of sample was loaded on Superdex$^{TM}$ 200 Increase 3.2/300 column (GE Healthcare Lifesciences) pre-equilibrated with buffer containing 25 mM Tris pH 8.0, 300 mM NaCl, 2 mM DTT. The fractions corresponding to each peak were collected and further analysed on a 4-12% SDS PAGE under reducing conditions followed by coomassie staining.

To check if UFL1/ UFBP1 could be reconstituted *in trans*, around 15 µgs of purified $His_6$-UFL1 and StrepII-UFBP1 were mixed and incubated on ice for 2h. Around 50 µls of sample was then loaded on Superdex$^{TM}$ 200 Increase 3.2/300 column (GE Healthcare Lifesciences) column pre-equilibrated with buffer containing 25 mM Tris pH 8.0, 300 mM NaCl, 2 mM DTT and the chromatograms were analysed.

**Preparation of UFC1$^{-O-UFM1}$**

UFC1$^{-O-UFM1}$ was prepared by incubating 40 µM UFC1$^{C116S}$ with 40 µM UFM1, 1 µM UBA5 in buffer containing 50 mM Tris pH 8.8, 10 mM $MgCl_2$, 5 mM ATP overnight at 25°C. The reaction was incubated briefly with 20 mM DTT to remove any non-specific di-sulfide linked adducts formed and passed through HiLoad$^{TM}$ 16/60 Superdex$^{TM}$ 75 pg with buffer containing 25 mM Tris 8.0, 150 mM NaCl, 5% Glycerol, 2 mM DTT. The fractions corresponding to UFC1$^{-O-UFM1}$ were collected and analysed on 4-12% SDS PAGE gel. Finally, the pure fractions containing UFC1$^{-O-UFM1}$ were pooled, concentrated and stored in -80°C until further use.

**Alphafold Predictions**

The structure of UFL1/UFBP1 was predicted using the ColabFold Google Colab notebook "AlphaFold2 using MMseqs2" [49] (https://github.com/sokrypton/ColabFold). The predicted model with the highest IDDT score is shown in the main figure and the PAE scores for different models ranking 1 to 5 are shown in **Fig S4A**.

**DALI Analysis**

DALI server was used to perform structural similarity analysis [69]. WH1 domain of UFL1(52-115) was used as the query model and searched against PDB25 database. The top ten models in the order of the Z-score is shown in **Fig S4B**.

**Multiple sequence alignments**

The sequences of target proteins and their homologs were manually downloaded from UNIPROT database in fasta format. Multiple sequence alignment was performed using Jalview version 2.11.1.7 programme (MAFFT algorithm module using L-INS-I) [70,71].

**Mass Photometry Data Acquisition and Analysis**

Protein samples were prepared in a solution containing 25 mM Tris pH 8.0, 300 mM NaCl, 2 mM DTT and stored on ice before loading onto the Refeyn One$^{MP}$ instrument (Refeyn). Typically, a set of protein standards (NativeMark™ Unstained Protein Standard, Invitrogen) are used for calibration and to generate a standard curve. Following calibration, approximately 10 µl of diluted protein sample in the concentration range of 5–25 nM was introduced into the flow-chamber and movies of 60s duration were recorded. Data acquisition was performed using Acquire MP (Refeyn Ltd, v1.1.3) and analysed using Discover MP software. Data shown here is representation of at least 3 independent acquisitions ($n \geq 3$).

**SEC MALS data acquisition and analysis**

Size exclusion chromatography and multi angle light scattering (SEC–MALS) experiments were performed on a Dionex Ultimate 3000 HPLC system with an inline Wyatt miniDAWN TREOS MALS detector and Optilab T-rEX refractive index detector. In addition, the elution profile of the protein was monitored with UV 280 attached to the Dionex system. For size exclusion chromatography, Superdex 200 10/300 gl column (GE Healthcare LifeSciences) was used. 50 µl of the purified UFL1/UFBP1 stored in buffer containing 25 mM Tris pH 8.0, 300 mM NaCl, 2 mM

DTT was loaded into the SEC column with Dionex auto loader at a concentration of 4 mg/ml and a flow rate of 0.3ml/min was maintained throughout the experiment. Molar masses spanning elution peaks were calculated using ASTRA v6.0.0.108 (Wyatt).

**Ribosome UFMylation assays**

Purified 60S ribosomes were a generous gift from the Puglisi lab, and purified as described previously [72,73]. *In vitro* UFMylation of ribosomes was performed at 30 ºC with 1 µM UFC1, 0.5 µM UBA5, 0.5 µM UFM1, 0 nM or 75 nM UFL1/UFBP1, 50 nM purified 60S ribosomes, and increasing concentrations of CDK5RAP3 (0, 38, 75, 150 or 375 nM) in a reaction buffer of 25 mM HEPES pH 7.5, and 100 mM NaCl. The reaction was quenched in 4 x SDS-Load buffer. Western blots show 10 min reaction time; immunoblots for RPL26 (Abcam, 59567) and UFL1 (Bethyl, A303-455M) were run on the same gel, which was cut and probed for these proteins separately. Immunoblots for CDK5RAP3 Bethyl, A300-871A) and UFM1 (Abcam, Ab109307) were run on separate gels.

**LC-MS/MS sample preparation, data acquisition and analysis**

First *in vitro* UFMylation reaction was performed to generate UFMylated products including free UFM1 chains. Then, the reaction products were run on a 4-12% SDS PAGE gel to separate the products based on electrophoretic mobility. Next, the bands corresponding to di-UFM1 chains were excised and in-gel digestion was performed [74]. Digested peptides were analysed by LC-MS/MS on an Exploris 480 coupled to an Ultimate 3000 nanoLC system (Thermo Fisher Scientific) for figure 2D and on an Exploris 240 (Thermo Fisher Scientific) coupled to an Evosep One (Evosep) for figure 4D. For the analysis performed on the Ultimate 3000, samples were loaded on a 100 µm x 2 cm trap column (Thermo Fisher Scientific #164564-CMD) and analysed on a 75 µm x 50 cm analytical column (Thermo Fisher Scientific #ES903) using a gradient from 3% to 35% Buffer B (80% LCMS grade acetonitrile, 0.08% formic acid in water) over 53 minutes. The columns were then washed with 95% Buffer B for 2 minutes prior equilibration in 97% Buffer A (0.1% formic acid in LCMS grade water). For the analysis performed on the Evosep One, samples were loaded onto the Evotips as per manufacturer recommendations and analysed using the 30 SPD Method.Peptides were then analysed in on either the Exploris 240 or 480 using data dependant with a MS1 resolution of 60000, AGC target of 300% and maximum injection time of 25 or 28 ms. Peptide were then fragmented using TOP 2

s method, MS2 resolution of 15000, NCE of 30 or 32%, AGC of 100% and maximum injection time of 100 ms. Peptide identification was performed in Mascot using a restricted and frequently updated database containing ~2000 protein sequences of interest (MRC db). Carbamidomethylation (C) was set at fixed modification and Oxidation (M), Dioxidation (M) and the addition of the dipeptide Glycine-Valine (K) were set as variable modifications. Peptides were searchd using an MS1 tolerance of 10 ppm and MS2 tolerance of 0.06 Da and a maximum of 2 missed cleavages were allowed. Only hits indentified with a FDR (False Discovery Rate) <1% were selected and further analysed in Scaffold viewer V5. For semi-quantitative analysis of VG modification on individual sites, Total Ion chromatogram (TIC) values obtained from the LC/MS and represented graphically.

## References:


1. Van Der Veen, A. G. & Ploegh, H. L. Ubiquitin-like proteins. *Annu. Rev. Biochem.* **81**, 323–357 (2012).
2. Sasakawa, H. *et al.* Solution structure and dynamics of Ufm1, a ubiquitin-fold modifier 1. *Biochem. Biophys. Res. Commun.* **343**, (2006).
3. Komatsu, M. *et al.* A novel protein-conjugating system for Ufm1, a ubiquitin-fold modifier. *EMBO J.* **23**, 1977–1986 (2004).
4. Cappadocia, L. & Lima, C. D. Ubiquitin-like Protein Conjugation: Structures, Chemistry, and Mechanism. *Chemical Reviews* vol. 118 889–918 (2018).
5. Walczak, C. P. *et al.* Ribosomal protein RPL26 is the principal target of UFMylation. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 1299–1308 (2019).
6. Wang, L. *et al.* UFMylation of RPL26 links translocation-associated quality control to endoplasmic reticulum protein homeostasis. *Cell Res.* **30**, 5–20 (2020).
7. Liang, J. R. *et al.* A Genome-wide ER-phagy Screen Highlights Key Roles of Mitochondrial Metabolism and ER-Resident UFMylation. *Cell* **180**, 1160-1177.e20 (2020).
8. Wang, Z. *et al.* MRE11 UFMylation promotes ATM activation. *Nucleic Acids Res.* **47**, (2019).
9. Liu, J. *et al.* UFMylation maintains tumour suppressor p53 stability by antagonizing its ubiquitination. *Nat. Cell Biol.* **22**, 1056–1063 (2020).
10. Qin, B. *et al.* UFL1 promotes histone H4 ufmylation and ATM activation. *Nat. Commun.* **10**, (2019).
11. Lee, L. *et al.* UFMylation of MRE11 is essential for maintenance of telomere length and hematopoietic stem cell survival. (2019) doi:10.1101/846477.
12. Yoo, H. M. *et al.* Modification of asc1 by ufm1 is crucial for erα transactivation and breast cancer development. *Mol. Cell* **56**, 261–274 (2014).
13. Gak, I. A. *et al.* UFMylation regulates translational homeostasis and cell cycle progression. *bioRxiv* (2020) doi:10.1101/2020.02.03.931196.
14. Xie, Z., Fang, Z. & Pan, Z. Ufl1/RCAD, a Ufm1 E3 ligase, has an intricate connection with ER stress. *International Journal of Biological Macromolecules* vol. 135 760–767 (2019).
15. Liu, J. *et al.* A critical role of DDRGK1 in endoplasmic reticulum homoeostasis via regulation of IRE1α stability. *Nat. Commun.* **8**, (2017).
16. Duan, R. *et al.* UBA5 mutations cause a new form of autosomal recessive cerebellar ataxia. *PLoS One* **11**, 1–12 (2016).
17. Nahorski, M. S. *et al.* Biallelic UFM1 and UFC1 mutations expand the essential role of ufmylation in brain development. *Brain* **141**, 1934–1945 (2018).
18. Tatsumi, K. *et al.* The Ufm1-activating enzyme Uba5 is indispensable for erythroid differentiation in mice. *Nat. Commun.* **2**, (2011).
19. Lemaire, K. *et al.* Ubiquitin fold modifier 1 (UFM1) and its target UFBP1 protect pancreatic beta cells from ER stress-induced apoptosis. *PLoS One* **6**, (2011).
20. Cai, Y. *et al.* UFBP1, a Key Component of the Ufm1 Conjugation System, Is Essential for Ufmylation-Mediated Regulation of Erythroid Development. *PLoS Genet.* **11**, (2015).
21. Cai, Y., Singh, N. & Li, H. Essential role of Ufm1 conjugation in the hematopoietic system. *Experimental Hematology* vol. 44 (2016).
22. Sung, H. K. *et al.* Two novel ubiquitin-fold modifier 1 (Ufm1)-specific proteases, UfSP1 and UfSP2. *J. Biol. Chem.* **282**, 5256–5262 (2007).
23. Oweis, W. *et al.* Trans-Binding Mechanism of Ubiquitin-like Protein Activation Revealed by a UBA5-UFM1 Complex. *Cell Rep.* **16**, 3113–3120 (2016).

24. Banerjee, S., Kumar, M. & Wiener, R. Decrypting UFMylation: How proteins are modified with UFM1. *Biomolecules* **10**, 1–14 (2020).
25. Mizushima, T. *et al.* Crystal structure of Ufc1, the Ufm1-conjugating enzyme. *Biochem. Biophys. Res. Commun.* **362**, 1079–1084 (2007).
26. Liu, G. *et al.* NMR and X-RAY structures of human E2-like ubiquitin-fold modifier conjugating enzyme 1 (UFC1) reveal structural and functional conservation in the metazoan UFM1-UBA5-UFC1 ubiquination pathway. *J. Struct. Funct. Genomics* **10**, 127–136 (2009).
27. Stewart, M. D., Ritterhoff, T., Klevit, R. E. & Brzovic, P. S. E2 enzymes: More than just middle men. *Cell Res.* **26**, 423–440 (2016).
28. Brzovic, P. S., Lissounov, A., Christensen, D. E., Hoyt, D. W. & Klevit, R. E. A UbcH5/ubiquitin noncovalent complex is required for processive BRCA1-directed ubiquitination. *Mol. Cell* **21**, 873–880 (2006).
29. Tatsumi, K. *et al.* A novel type of E3 ligase for the Ufm1 conjugation system. *J. Biol. Chem.* **285**, 5417–5427 (2010).
30. Cai, Y. *et al.* Indispensable role of the Ubiquitin-fold modifier 1-specific E3 ligase in maintaining intestinal homeostasis and controlling gut inflammation. *Cell Discov.* **5**, (2019).
31. Deol, K. K., Lorenz, S. & Strieter, E. R. Enzymatic Logic of Ubiquitin Chain Assembly. *Front. Physiol.* **10**, 1–14 (2019).
32. Pichler, A., Fatouros, C., Lee, H. & Eisenhardt, N. SUMO conjugation - A mechanistic view. *Biomolecular Concepts* vol. 8 13–36 (2017).
33. Lorenz, S. Structural mechanisms of HECT-type ubiquitin ligases. *Biol. Chem.* **399**, 127–145 (2018).
34. Deshaies, R. J. & Joazeiro, C. A. P. RING domain E3 ubiquitin ligases. *Annu. Rev. Biochem.* **78**, 399–434 (2009).
35. Walden, H. & Rittinger, K. RBR ligase–mediated ubiquitin transfer: a tale with many twists and turns. *Nat. Struct. Mol. Biol.* **25**, 440–445 (2018).
36. Pao, K. *et al.* Ligase With Esterification Activity. *Nature* (2018).
37. Wu, J., Lei, G., Mei, M., Tang, Y. & Li, H. A novel C53/LZAP-interacting protein regulates stability of C53/LZAP and DDRGK domain-containing protein 1 (DDRGK1) and modulates NF-κB signaling. *J. Biol. Chem.* **285**, 15126–15136 (2010).
38. Witting, K. F. & Mulder, M. P. C. Highly specialized ubiquitin-like modifications: Shedding light into the ufm1 enigma. *Biomolecules* vol. 11 (2021).
39. Gerakis, Y., Quintero, M., Li, H. & Hetz, C. The UFMylation System in Proteostasis and Beyond. *Trends in Cell Biology* vol. 29 (2019).
40. Kumar, M. *et al.* Structural basis for UFM1 transfer from UBA5 to UFC1. *Nat. Commun.* **12**, (2021).
41. Padala, P. *et al.* Novel insights into the interaction of UBA5 with UFM1 via a UFM1-interacting sequence. *Sci. Rep.* **7**, 1–12 (2017).
42. Soudah, N. *et al.* An N-Terminal Extension to UBA5 Adenylation Domain Boosts UFM1 Activation: Isoform-Specific Differences in Ubiquitin-like Protein Activation. *J. Mol. Biol.* **431**, 463–478 (2019).
43. Wenzel, D. M., Lissounov, A., Brzovic, P. S. & Klevit, R. E. UBCH7 reactivity profile reveals parkin and HHARI to be RING/HECT hybrids. *Nature* **474**, 105–108 (2011).
44. Kihm, A. J. *et al.* An abundant erythroid protein that stabilizes free α-haemoglobin. *Nature* **417**, 758–763 (2002).
45. Warner, J. R. In the absence of ribosomal RNA synthesis, the ribosomal proteins of HeLa Cells are synthesized normally and degraded rapidly. *J. Mol. Biol.* **115**, 315–333 (1977).
46. Sonn-Segev, A. *et al.* Quantifying the heterogeneity of macromolecular machines by

mass photometry. *Nat. Commun.* **11**, 1–10 (2020).

47. Deng, L. *et al. Activation of the IB Kinase Complex by TRAF6 Requires a Dimeric Ubiquitin-Conjugating Enzyme Complex and a Unique Polyubiquitin Chain through covalent modification by Ub. Cell* vol. 103 (2000).
48. Branigan, E., Plechanovová, A., Jaffray, E. G., Naismith, J. H. & Hay, R. T. Structural basis for the RING-catalyzed synthesis of K63-linked ubiquitin chains. *Nat. Struct. Mol. Biol.* **22**, 597–602 (2015).
49. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
50. Stewart, M. Structural biology of the PCI-protein fold. *Bioarchitecture* vol. 2 118–123 (2012).
51. Scheel, H. & Hofmann, K. Prediction of a common structural scaffold for proteasome lid, COP9-signalosome and eIF3 complexes. *BMC Bioinformatics* **6**, (2005).
52. Yang, R. *et al.* CDK5RAP3, a UFL1 substrate adaptor, is crucial for liver development. *Dev.* **146**, (2019).
53. Pichler, A., Knipscheer, P., Saitoh, H., Sixma, T. K. & Melchior, F. The RanBP2 SUMO E3 ligase is neither HECT- nor RING-type. *Nat. Struct. Mol. Biol.* **11**, 984–991 (2004).
54. Plechanovov, A., Jaffray, E. G., Tatham, M. H., Naismith, J. H. & Hay, R. T. Structure of a RING E3 ligase and ubiquitin-loaded E2 primed for catalysis. *Nature* **489**, 115–120 (2012).
55. Pruneda, J. N. *et al.* Structure of an E3:E2~Ub Complex Reveals an Allosteric Mechanism Shared among RING/U-box Ligases. *Mol. Cell* **47**, 933–942 (2012).
56. Reverter, D. & Lima, C. D. Insights into E3 ligase activity revealed by a SUMO–RanGAP1–Ubc9–Nup358 complex. *Nature* **435**, 687–692 (2005).
57. Cappadocia, L., Pichler, A. & Lima, C. D. Structural basis for catalytic activation by the human ZNF451 SUMO E3 ligase. *Nat. Struct. Mol. Biol.* **22**, 968–975 (2015).
58. Gundogdu, M. & Walden, H. Structural basis of generic versus specific E2–RING E3 interactions in protein ubiquitination. *Protein Sci.* **28**, 1758–1770 (2019).
59. Middleton, A. J., Wright, J. D. & Day, C. L. Regulation of E2s: A Role for Additional Ubiquitin Binding Sites? *J. Mol. Biol.* **429**, 3430–3440 (2017).
60. Schumacher, F.-R., Wilson, G. & Day, C. L. The N-Terminal Extension of UBE2E Ubiquitin-Conjugating Enzymes Limits Chain Assembly. *J. Mol. Biol.* **425**, 4099–4111 (2013).
61. Brzovic, P. S. *et al.* Binding and recognition in the assembly of an active BRCA1/BARD1 ubiquitin-ligase complex. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 5646–5651 (2003).
62. Brown, N. G. *et al.* RING E3 mechanism for ubiquitin ligation to a disordered substrate visualized for human anaphase-promoting complex. *Proc. Natl. Acad. Sci.* **112**, 5272–5279 (2015).
63. Takahashi, T. S. *et al.* Structural basis of ubiquitin recognition by the winged-helix domain of Cockayne syndrome group B protein. *Nucleic Acids Res.* **47**, 3784–3794 (2019).
64. Harami, G. M., Gyimesi, M. & Kovács, M. From keys to bulldozers: Expanding roles for winged helix domains in nucleic-acid-binding proteins. *Trends in Biochemical Sciences* vol. 38 364–371 (2013).
65. ARAVIND, L., ANANTHARAMAN, V., BALAJI, S., BABU, M. & IYER, L. The many faces of the helix-turn-helix domain: Transcription regulation and beyond. *FEMS Microbiol. Rev.* **29**, 231–262 (2005).
66. Qin, B. *et al.* STK38 promotes ATM activation by acting as a reader of histone H4 ufmylation. *Sci. Adv.* **6**, 1–10 (2020).

67. Stephani, M. *et al.* A cross-kingdom conserved er-phagy receptor maintains endoplasmic reticulum homeostasis during stress. *Elife* **9**, (2020).

68. Baek, K., Scott, D. C. & Schulman, B. A. NEDD8 and ubiquitin ligation by cullin-RING E3 ligases. *Curr. Opin. Struct. Biol.* **67**, 101–109 (2021).

69. Holm, L. & Rosenström, P. Dali server: Conservation mapping in 3D. *Nucleic Acids Res.* **38**, (2010).

70. Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M. & Barton, G. J. Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189–1191 (2009).

71. Katoh, K. & Toh, H. Parallelization of the MAFFT multiple sequence alignment program. *Bioinformatics* **26**, 1899–1900 (2010).

72. Johnson, A. G. *et al.* RACK1 on and off the ribosome. *RNA* **25**, 881–895 (2019).

73. Lapointe, C. P. *et al.* Dynamic competition between SARS-CoV-2 NSP1 and mRNA on the human ribosome inhibits translation initiation. *Proc. Natl. Acad. Sci.* **118**, e2017715118 (2021).

74. Shevchenko, A., Tomas, H., Havliš, J., Olsen, J. V. & Mann, M. In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat. Protoc.* **1**, 2856–2860 (2007).

## Figures

### Figure 1

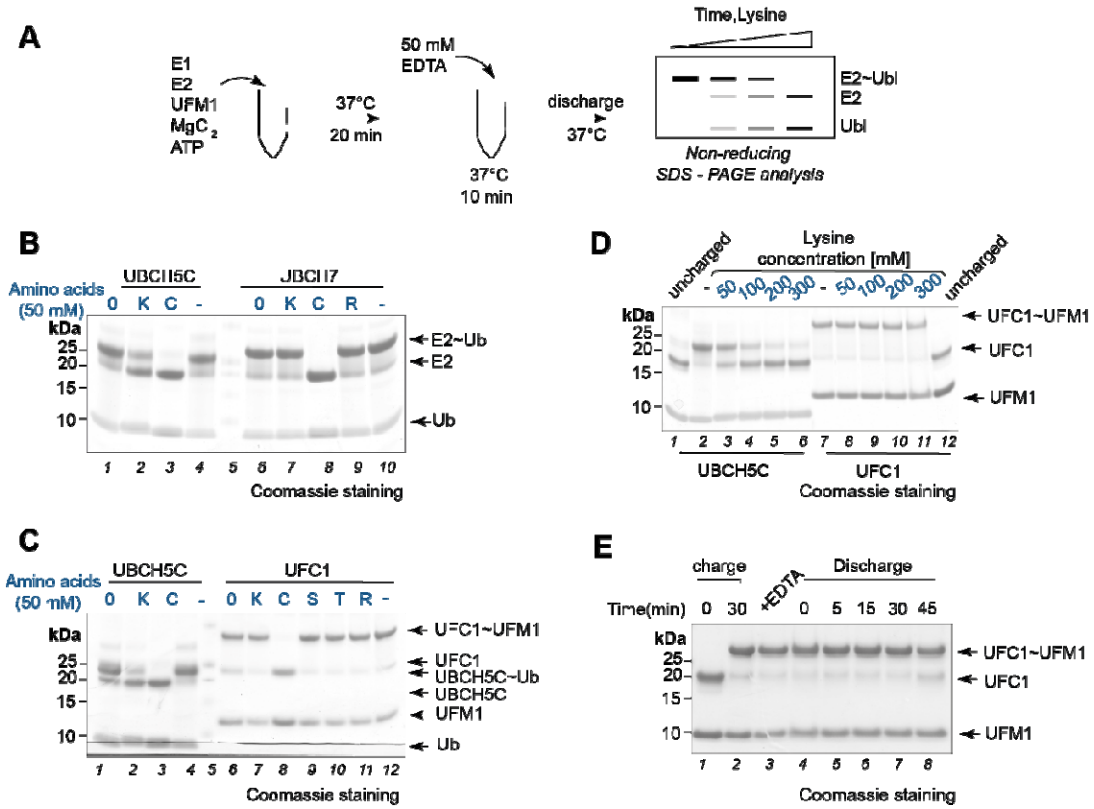



**Figure 1: - Nucleophilic reactivity profiling of UFC1**

- **A.** Schematic showing the workflow for substrate independent single turnover lysine discharge assays.
- **B.** Assay to monitor discharge of Ubiquitin from UBCH5C and UBCH7 in the presence of free amino acids. 0 indicates time 0.
- **C.** Assay to check for discharge of UFM1 from UFC1 in the presence of free amino acids. UBCH5C is used as a positive control for lysine and cysteine discharge.
- **D.** Lysine discharge assays using UBCH5C and UFC1 in the presence of increasing concentration of free Lysine.
- **E.** Time-dependent analysis of discharge of UFM1 from UFC1 in the presence of high concentration of free lysine (150 mM).

## Figure 2

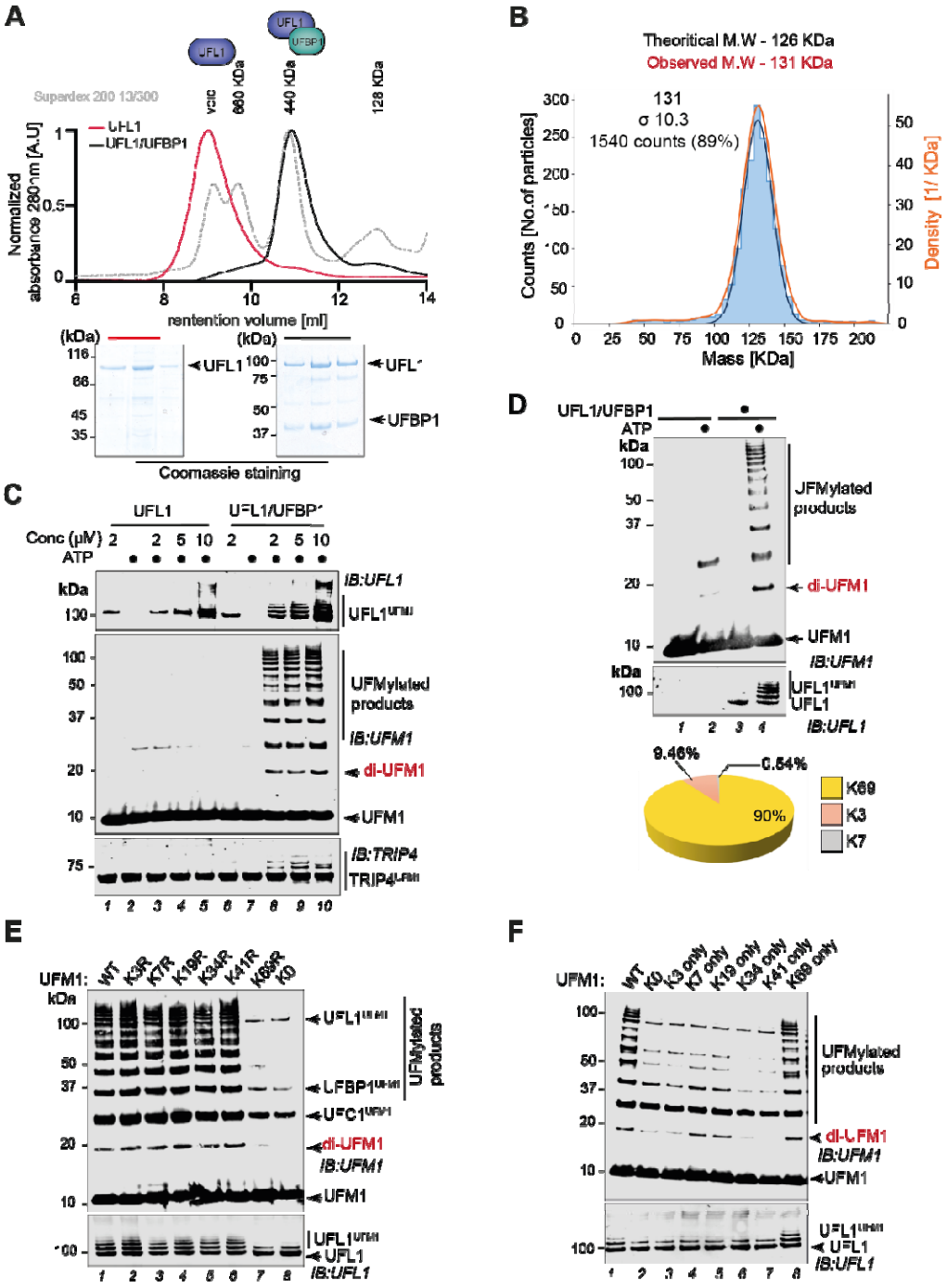



**Figure 2: - *In vitro* reconstitution of the active UFM1 E3 ligase.**

**A.** Comparison of size exclusion chromatography profiles of UFL1 expressed alone (Red) and co-expressed UFL1/UFBP1 complex (Black) run under identical buffer conditions on a HiLoad$^{TM}$ Superdex 10/300 gl column. Molecular weight standards are shown in grey. Fractions corresponding to

each peak were collected and analysed on 4-12% denaturing SDS-PAGE which is shown below.

**B.** Mass photometry analysis of co-purified UFL1/UFBP1 complex. The theoretical and experimental molecular weights are indicated above.

**C.** Immunoblot comparing UFL1 autoUFMylation *(Top)* and UFM1 chain synthesis *(Middle)* in the presence of UFL1 expressed alone and in complex with UFBP1. *(Bottom)* Assays to monitor UFMylation of substrates using purified TRIP4 in the presence of UFL1 alone and UFL1/UFBP1 complex. Reaction products were run on a 4-12% SDS-PAGE gel and analysed by immunoblotting using indicated antibodies.

**D.** Immunoblot showing formation of free UFM1 chains in the presence and absence of UFL1/UFBP1. *(Bottom)* Schematic representation of the linkage composition of di-UFM1 chains obtained from LC-MS/MS analysis.

**E.** *In vitro* UFMylation assay in the presence of Lys to Arg (K-R) and lysine less(K0) mutants of UFM1 to check for polyUFMylation *(top)* and UFL1 autoUFMylation *(bottom)* .

**F.** *In vitro* UFMylation assay to check for formation of di-UFM1 chains in the presence of single Lys and Lys less mutants of UFM1.

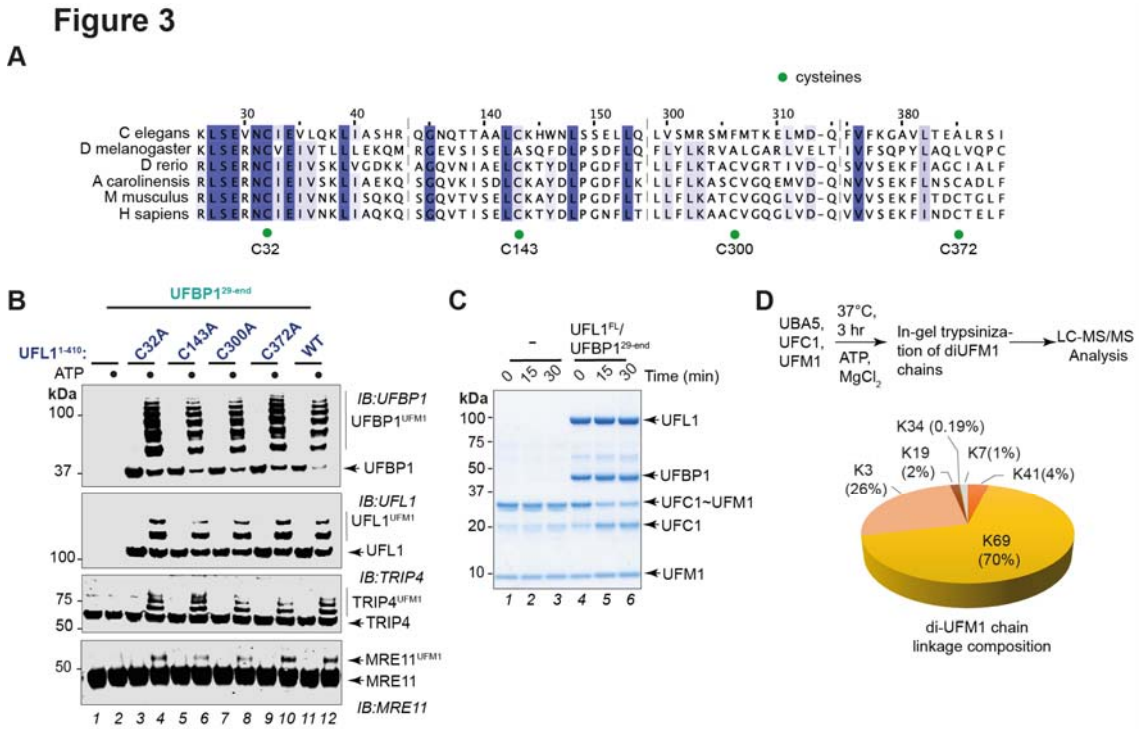


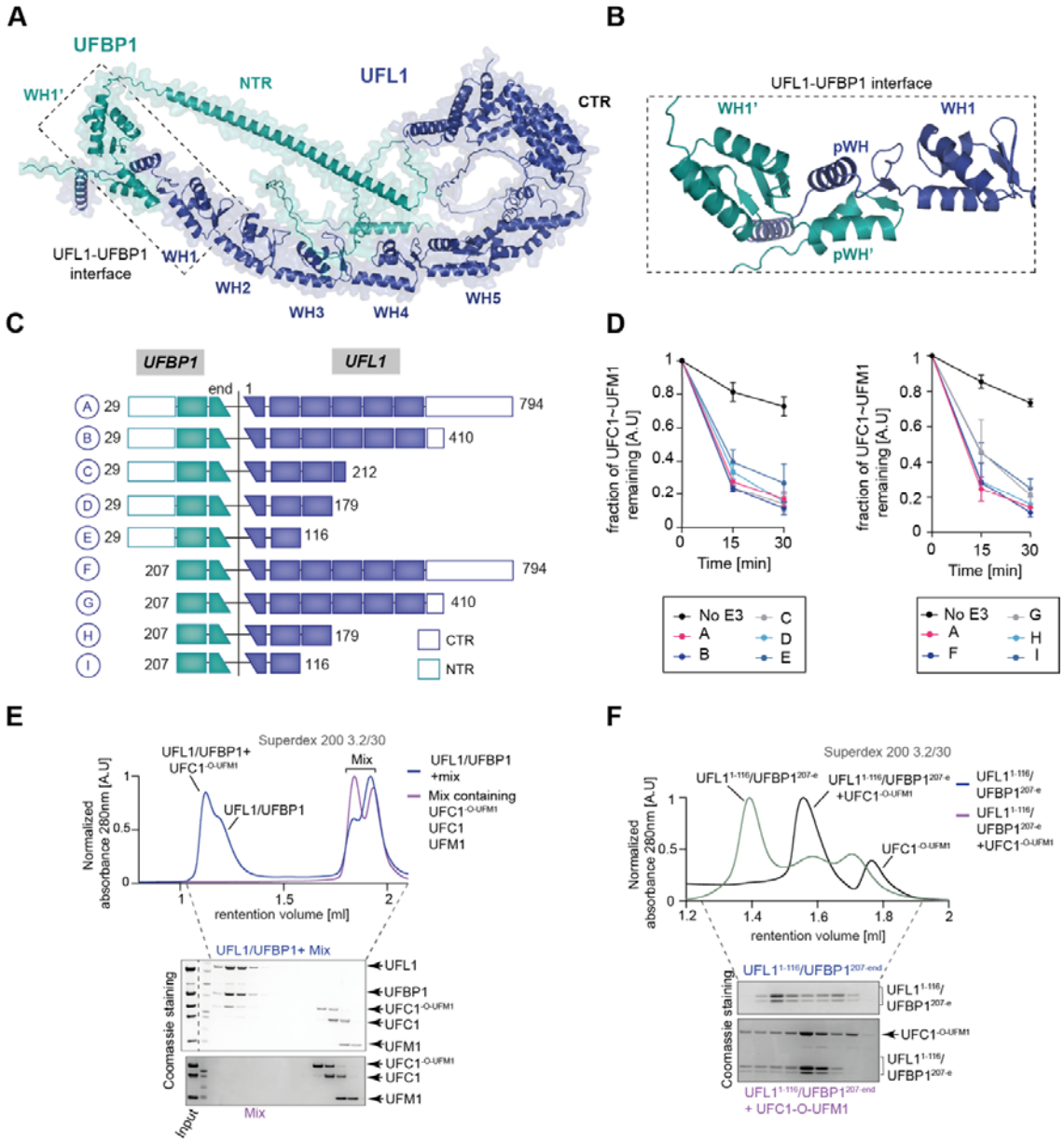

**Figure 3: - UFL1/UFBP1 is a scaffold type E3 ligase**

**A.** Multiple sequence alignment of 1-410 a.a. region of UFL1 from various organisms to highlight the degree of conservation of cysteine residues in this region.

**B.** Immunoblot showing *in vitro* UFMylation assays in the presence of C to A mutants of UFL1 to check for UFBP1 UFMylation *(top)*, UFL1 autoUFMylation *(middle)* and substrate UFMylation *(bottom)*.

**C.** Coomassie stained SDS-PAGE gel showing discharge assays in the presence and absence of full length UFL1/UFBP1 complex. Charged UFC1 was incubated with lysine (50 mM ) in the presence and absence of full length UFL1/UFBP1 complex and analysed for discharge on a 4-12% SDS PAGE gel under non-reducing conditions.

**D.** Graphical representation showing the composition of linkage forms of di-UFM1 chains formed by minimal reconstitution of UBA5 and UFC1.




991
992
993
994

**Figure 4: - UFL1/UFBP1 complex functions as a scaffold type E3 ligase**

**A.** Cartoon representation of predicted full length human UFL1/UFBP1 complex using Alphafold. The individual winged helix domains of UFL1 and UFBP1 is shown in blue and cyan respectively. The dimeric interface of UFL1 and UFBP1 is highlighted in the dotted box.

**B.** Enlarged view of the dimeric interface of UFL1/UFBP1 complex shown to highlight the formation of full WH domain bridged by two partial winged helix domains of UFL1 (shown in blue) and UFBP1 (shown in teal).

**C.** Schematic representation of UFL1/UFBP1 constructs with different domain boundaries to identify the minimal catalytic region of UFL1/UFBP1 complex required for aminolysis.

**D.** Quantitative representation of Lysine discharge assays to identify the minimal boundaries of UFL1 and UFBP1 required for activation of UFC1 (n=3, mean±SDs.

**E.** Size exclusion chromatography profiles showing that full length UFL1/UFBP1 complex interacts specifically with charged E2. UFL1/UFBP1 complex was incubated with UFC1$^{-O\text{-}UFM1}$ at the molar ratio of 1:1 for 20 min at 4°C and loaded on a Superdex 200 Increase 3.2/300 column. The fractions corresponding to each peak were collected and analysed on a 4-12% SDS PAGE gel followed by Coomassie staining.

**F.** Minimal catalytic region is sufficient for interaction with charged UFC1. UFL1$^{(1\text{-}179)}$/UFBP1$^{(1\text{-}116)}$ complex was incubated with UFC1$^{-O\text{-}UFM1}$ at the molar ratio of 1:1 for 20 min at 4°C and analysed by analytical size exclusion chromatography as described in E).

**G.** *In vitro* UFMylation assay to monitor formation of free di-UFM1 chains in the presence of UFL1/UFBP1 complexes bearing different domain boundaries.

**H.** Substrate UFMylations assays to check for UFMylation of purified Histone H4 in the presence of different UFL1/UFBP1 truncations.

**I.** Role of UFBP1 in substrate UFMylation. *(Top)*Comparison of UFMylation activities of UFL1/UFBP1$^{29\text{-}end}$ and UFL1/UFBP1$^{207\text{-}end}$ using Histone H4. *(Bottom)*Comparison of UFMylation activities of UFL1$^{1\text{-}179}$/UFBP1$^{29\text{-}end}$ and UFL1$^{1\text{-}179}$/UFBP1$^{207\text{-}end}$ using Histone H4.

## Figure 5

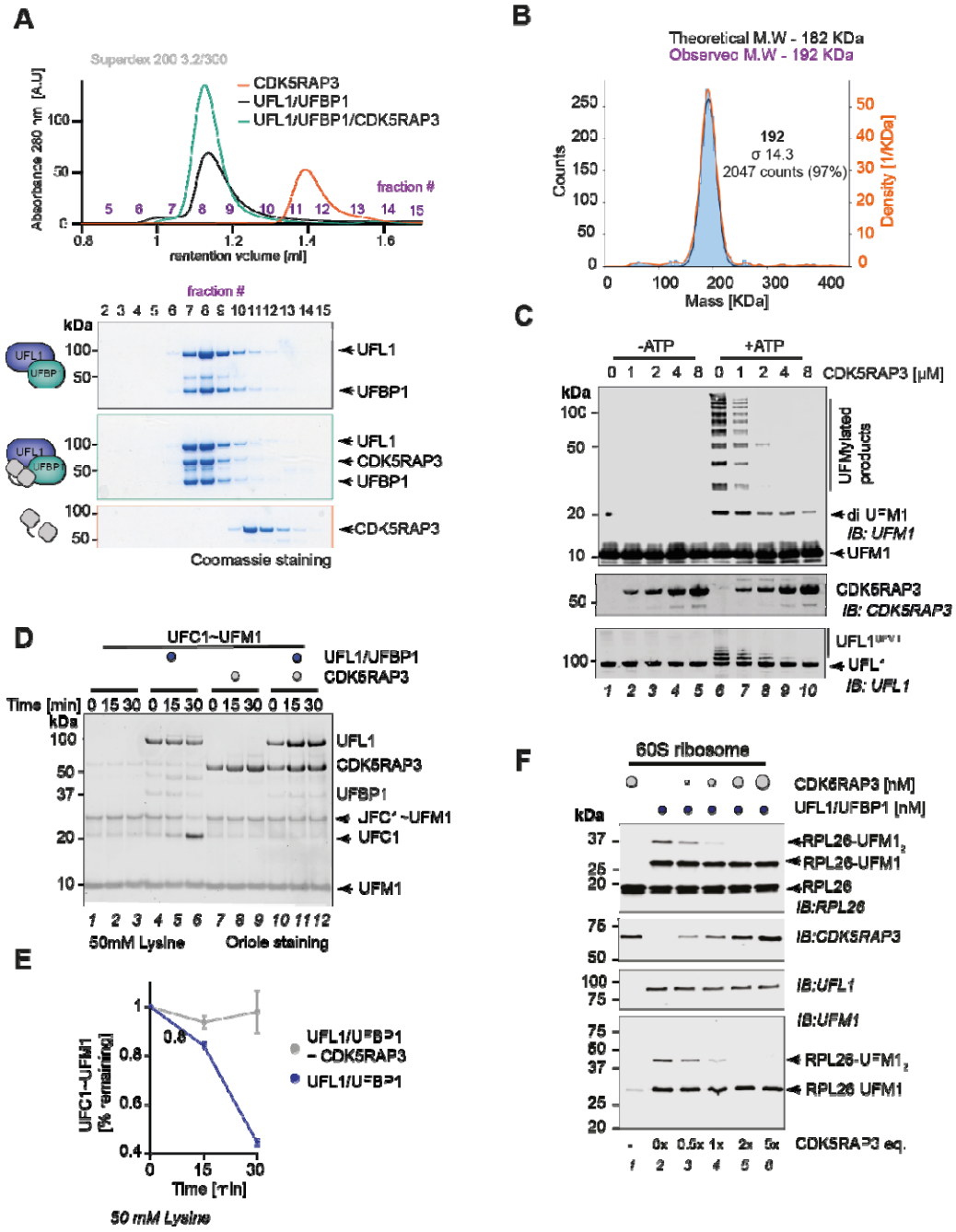



**Figure 5: - CDK5RAP3 inhibits E3 ligase activity of UFL1/UFBP1 *in vitro***

**A.** CDK5RAP3 forms a complex with UFL1/UFBP1 *in vitro*. 30 µg of UFL1/UFBP1 was mixed with 15 µg of CDK5RAP3 and loaded on an analytical gel filtration column. (Bottom) Fractions were collected and analysed on 4-12% SDS PAGE and visualized by Coomassie staining.

**B.** Mass photometry analysis showing the experimental molecular weight of UFL1/UFBP1/CDK5RAP3 complex.

**C.** *In vitro* UFMylation assay in the presence of increasing concentrations of CDK5RAP3 to monitor the E3 ligase activity of UFL1/UFBP1 complex. The reaction products were run on a 4-12% SDS-PAGE gel and immunoblotting was performed using indicated antibodies.

**D.** *In vitro* UFMylation assay to monitor the UFMylation of purified TRIP4 in the presence and absence of CDK5RAP3.

**E.** Lysine discharge assays to check for activation of UFC1 by UFL1/UFBP1 in the presence and absence of CDK5RAP3. The reaction products were run on a 4-12% SDS PAGE and visualized by Oriole staining.

**F.** Quantification of discharge of UFM1 from UFC1 in the presence and absence of CDK5RAP3 as seen in E). n=3, mean ±SDs.

## Figure 6

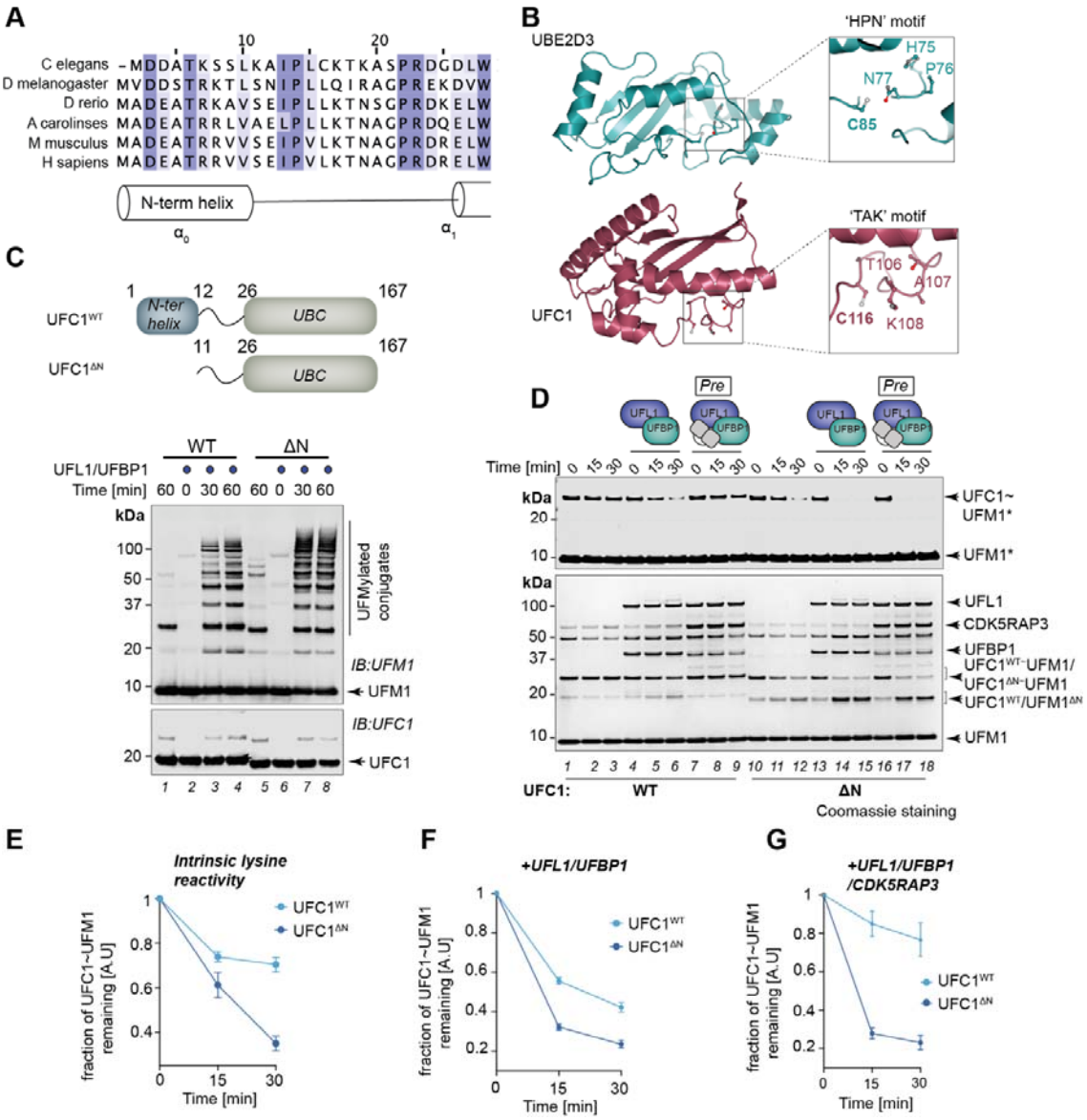



**Figure 6:- Role of the N-terminal helix of UFC1 in UFMylation**

**A.** Multiple sequence alignment of UFC1 homologs from various organisms. A graphical representation of the secondary structure is extracted from the crystal structure of UFC1 (PDB ID:2Z6O)

**B.** Crystal structures of UBE2D3 (PDB ID:5EGG) and UFC1 (PDB ID:2Z6O) are shown in cartoon representation. *(Inset)* Catalytic cysteines and critical residues required for catalytic activity of UBE2D3 (green) and UFC1(dark red) are labelled and shown in ball and stick representation.

**C.** *In vitro* UFMylation assays in the presence of Wild type UFC1 ($UFC1^{WT}$) and UFC1 lacking the N-terminal helix ($UFC1^{\Delta N}$). Schematic representation of domain features of $UFC1^{WT}$ and $UFC1^{\Delta N}$ is shown above.

**D.** Assay to compare Lys discharge activities of $UFC1^{WT}$ and $UFC1^{\Delta N}$ on its own, in the presence of UFL1/UFBP1 or in the presence of preassembled UFL1/UFBP1/CDK5RAP3 complex. Top gel: LiCOR scan of fluorescently labelled UFM1; bottom gel – Coomassie stained (Representative of three independent experiments).

**E.** Quantitative analysis of the intrinsic Lys reactivity (Lysine 25 mM) of $UFC1^{WT}$ and in $UFC1^{\Delta N}$ in the absence of E3 ligase.

**F.** Quantitative analysis of Lys reactivity of $UFC1^{WT}$and in $UFC1^{\Delta N}$ in the presence of UFL1/UFBP1.

**G.** Quantitative analysis of Lys reactivity of $UFC1^{WT}$and in $UFC1^{\Delta N}$ in the the presence of preassembled UFL1/UFBP1/CDK5RAP3. E – G: n=3, mean ±SD.

# Supplementary Material

## Figure S1

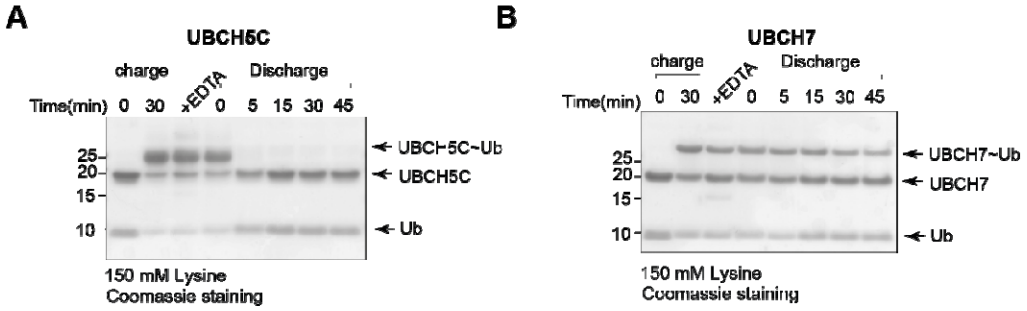



**Figure S1: Analysis of E2 discharge**

**A.** & **B.** Time-dependent analysis of discharge of Ubiquitin from UBCH5C and UBCH7 in the presence of 150 mM.

## Figure S2

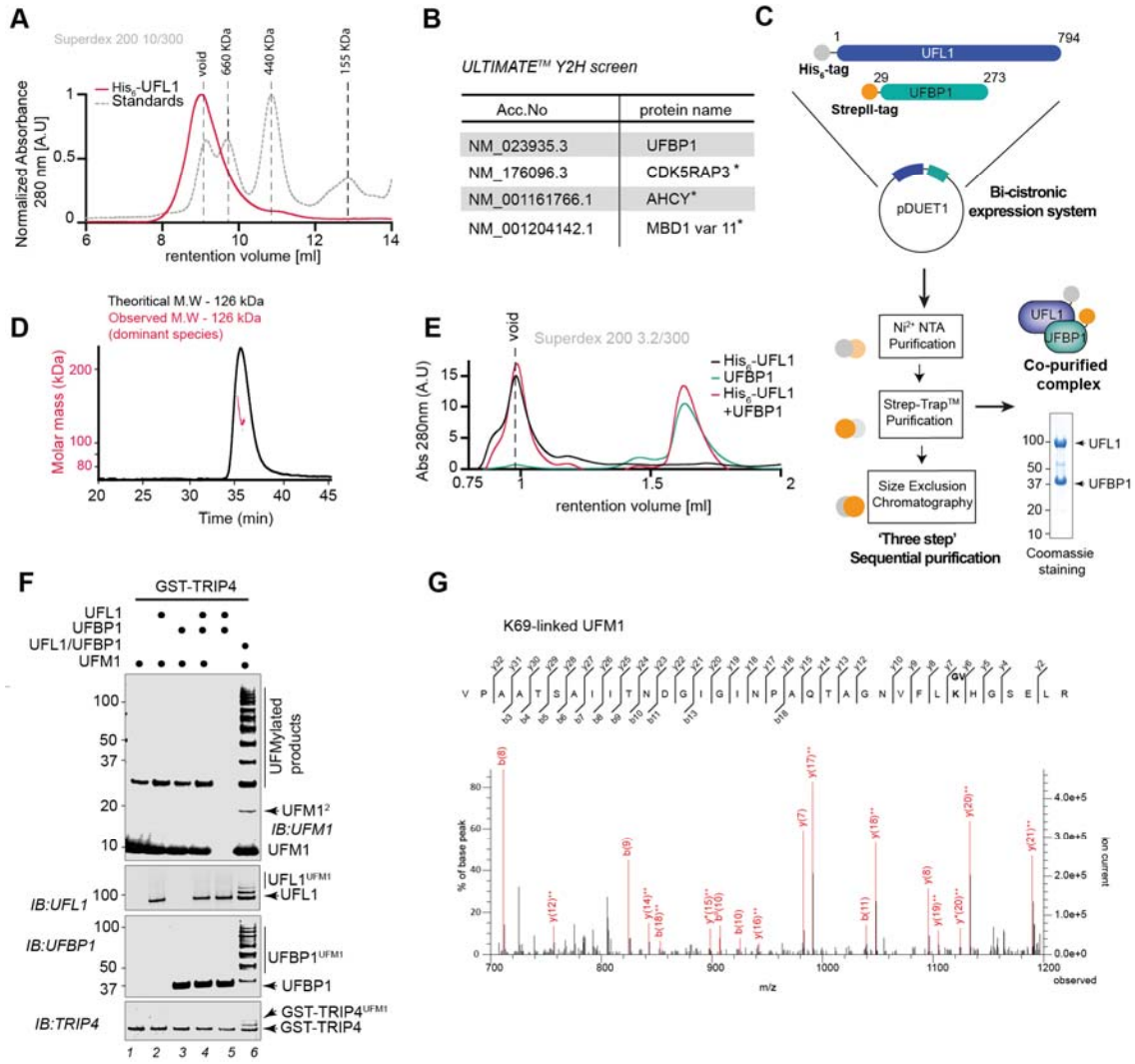



**Figure S2: UFL1 is only active in the presence of co-expressed UFBP1**

**A.** Size exclusion chromatography profile of UFL1 run on Superdex 200 Increase 10/300gl column (shown in red). Overlay of chromatogram of molecular weight standards of different sizes (shown in grey) run under same buffer conditions.

**B.** Results obtained from ULTIMATE Y2H$^{TM}$ screening (Hybrigenics) for binary interactions with UFL1. Asterisk denotes protein hits obtained with low confidence.

**C.** Schematic describing the strategy for co-expression and purification of UFL1/UFBP1 complex.

**D.** SEC-MALS analysis of UFL1/UFBP1 complex. The theoretical and observed molecular weights are indicated.

**E.** Analytical gel filtration chromatography analysis of UFL1/UFBP1 complex and its subunits run on Superdex 200 3.2/300.

**F.** UFBP1 does not activate UFL1 when added exogenously. *In vitro* UFMylation assays to compare the E3 ligase activity of UFL1 and UFBP1 expressed alone and together as a complex. 0.25 µM UBA5, 5 µM UFC1 and 10 µM UFM1 was incubated with 1 µM UFL1 or 1 µM UFBP1 or 1 µM UFL1/UFBP1 complex for 1 h at 37°C in the presence of 50 mM HEPES 7.5, 0.5 mM DTT, 10 mM $MgCl_2$ and 10 mM ATP. The reaction was stopped by addition of 3x SDS loading dye and run on a 4-12% denaturing SDS PAGE gel under reducing conditions and immunoblotting was performed using indicated antibodies.

**G.** $MS^2$ spectra showing the peptide derived from *in vitro* UFMylation assay showing the VG-remnant on K69 of UFM1.

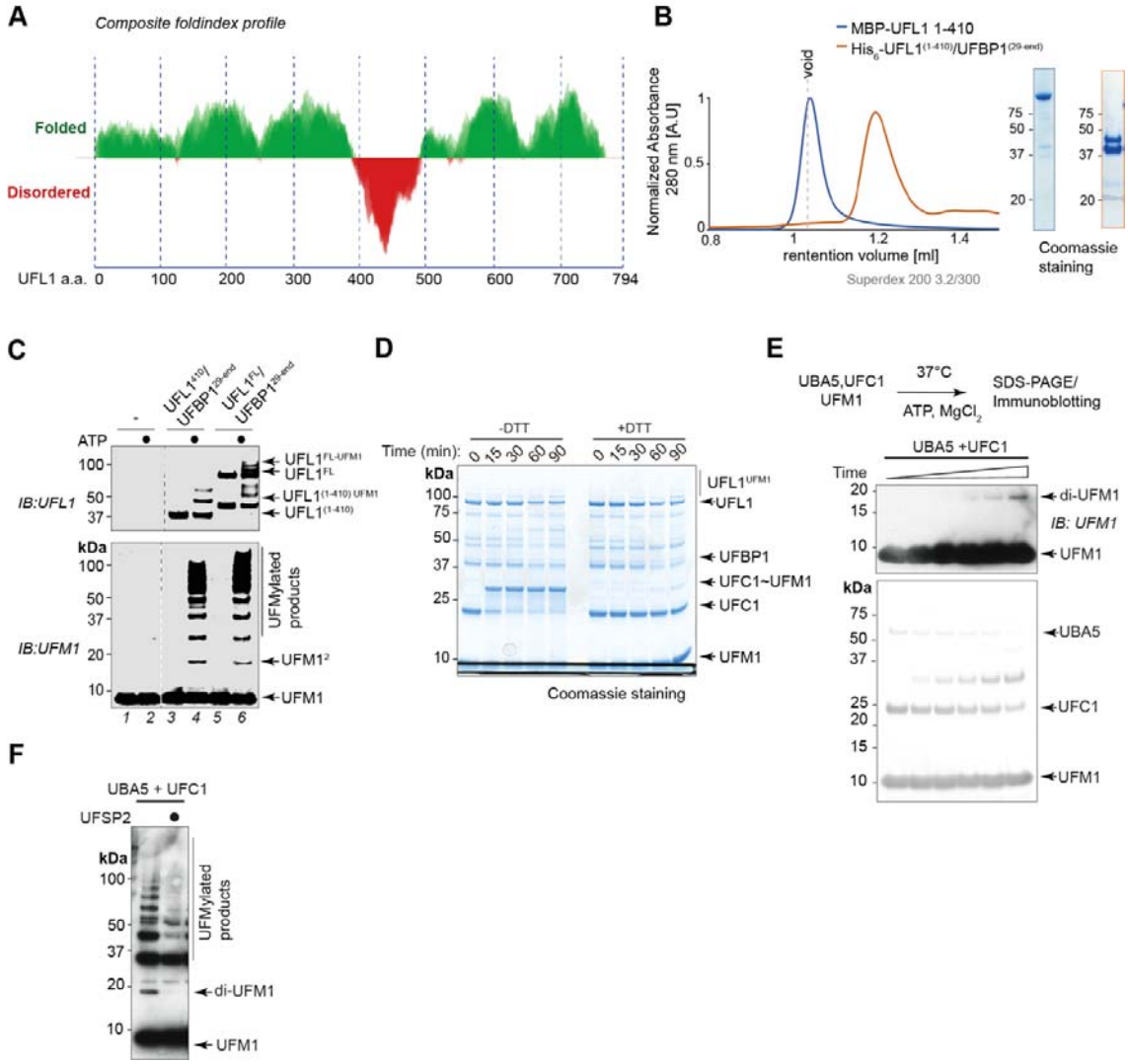



**Figure S3: Mode of UFMylation by UFL1/UFBP1 complex**

**A.** Composite Foldindex profile of UFL1 showing folding propensity of different regions of UFL1 to aid in construct design for soluble protein expression.

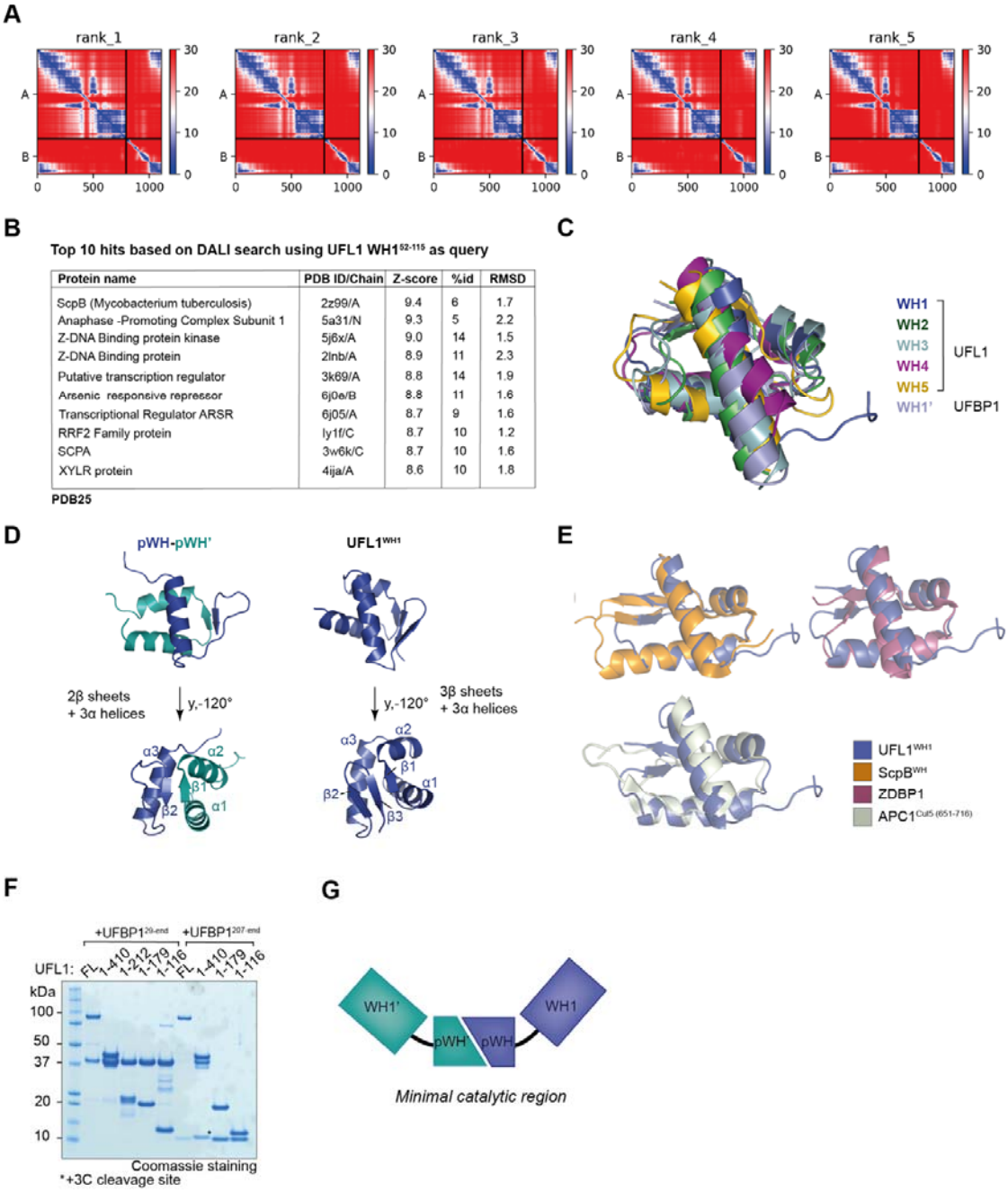
**B.** Size exclusion chromatography profile of MBP-UFL1$^{1-410}$ (dark blue) and His$_6$-UFL1$^{1-410}$/UFBP1$^{29-end}$(orange) run on Superdex 200 3.2/300 column (shown in red). Approximately 20 µgs of sample was run on a Superdex 200 3.2/300 analytical gel filtration column and UV traces measured at 280 nm were recorded and analysed. (Right) Coomassie stained gel showing the purity of the proteins.

**C.** *In vitro* UFMylation assay to check for E3 ligase activity of UFL1$^{1-410}$/UFBP1$^{29-end}$. Full length UFL1/UFBP1 complex is used as a positive control.

**D.** Time course assay to check for transthiolation activity of UFL1. Reaction products were analysed on a 4-12% SDS PAGE gel under reducing and non-reducing conditions.

**E.** *In vitro* UFMylation assays to check for formation of di-UFM1 chains in the by minimal reconstitution in the presence UBA5 and UFC1 alone.

**F.** Immunoblot to check for the presence of UFM1 chains with and without treatment of UFSP2.




| Protein name | PDB ID/Chain | Z-score | %id | RMSD |
|---|---|---|---|---|
| ScpB (Mycobacterium tuberculosis) | 2z99/A | 9.4 | 6 | 1.7 |
| Anaphase -Promoting Complex Subunit 1 | 5a31/N | 9.3 | 5 | 2.2 |
| Z-DNA Binding protein kinase | 5j6x/A | 9.0 | 14 | 1.5 |
| Z-DNA Binding protein | 2lnb/A | 8.9 | 11 | 2.3 |
| Putative transcription regulator | 3k69/A | 8.8 | 14 | 1.9 |
| Arsenic responsive repressor | 6j0e/B | 8.8 | 11 | 1.6 |
| Transcriptional Regulator ARSR | 6j05/A | 8.7 | 9 | 1.6 |
| RRF2 Family protein | 1y1f/C | 8.7 | 10 | 1.2 |
| SCPA | 3w6k/C | 8.7 | 10 | 1.6 |
| XYLR protein | 4ija/A | 8.6 | 10 | 1.8 |

**Figure S4: - Identifying the minimal ligase domain of UFL1/UFBP1 complex**

**A.** Predicted Aligned Error (PAE) scores of predicted UFL1/UFBP1 models using Alphafold.

**B.** List of structurally similar proteins from the PDB25 database predicted using UFL1 WH1(52-115) domain as query structure using DALI server.

**C.** Superposition of the winged helix (WH) domains of UFL1 and UFBP1.

**D.** Comparison of composite winged helix domain formed by two hemi WH domains of UFL1 (pWH) and UFBP1 (pWH') with the WH1 domain of UFL1.

**E.** Overlaid structures of Winged helix 1(WH1) domain of UFL1(52-115) and top three hits obtained from DALI search shown in cartoon representation.

**F.** Coomassie stained gel showing purity of different UFL1/UFBP1 truncations.

**G.** Schematic showing the tandem WH domains which constitute the minimal ligase domain.

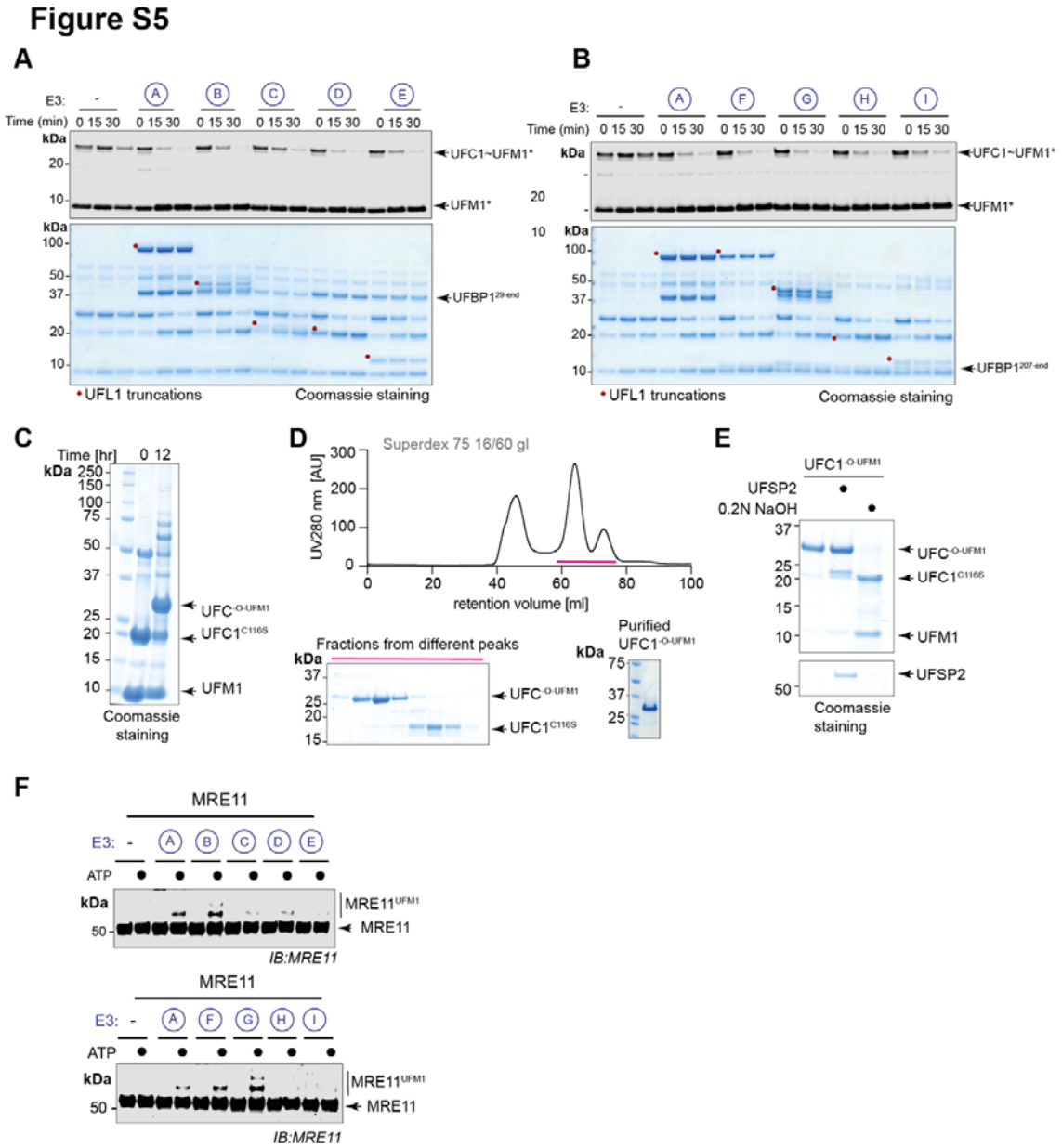



**Figure S5: - UFL1/UFBP1 E3 ligase complex binds to UFC1~UFM1 and induces aminolysis**

**A.** Lysine discharge assays to check for activation of UFC1 in the presence of different UFL1 truncations. Top gel: LiCOR scan of fluorescently labelled UFM1; bottom gel – Coomassie stained (Representative of three independent experiments).

**B.** Lysine discharge assays as in A) in the absence of NTR of UFBP1(representative of 3 independent experiments).

**C.** Coomassie stained gel showing analysis of *in vitro* UFMylation reaction products to check for the formation of stable oxy-ester linked UFC1-UFM1 conjugate ($UFC1^{-O-UFM1}$).

**D.** Chromatogram obtained from SEC analysis using HiLoad™ Superdex 75 16/60pg column. (Bottom left) The fractions collected were run on an SDS-PAGE gel to identify fractions that contained pure $UFC1^{-O-UFM1}$. (Bottom right) Coomassie stained gel showing analysis of purified $UFC1^{-O-UFM1}$ product to check for homogeneity.

**E.** Quality check to analyse if UFC1-UFM1 conjugate is linked through an oxy-ester linkage by alkaline hydrolysis.

**F.** Substrate UFMylations assays to check for UFMylation of MRE11 in the presence of different UFL1/UFBP1 truncations.

## Figure S6

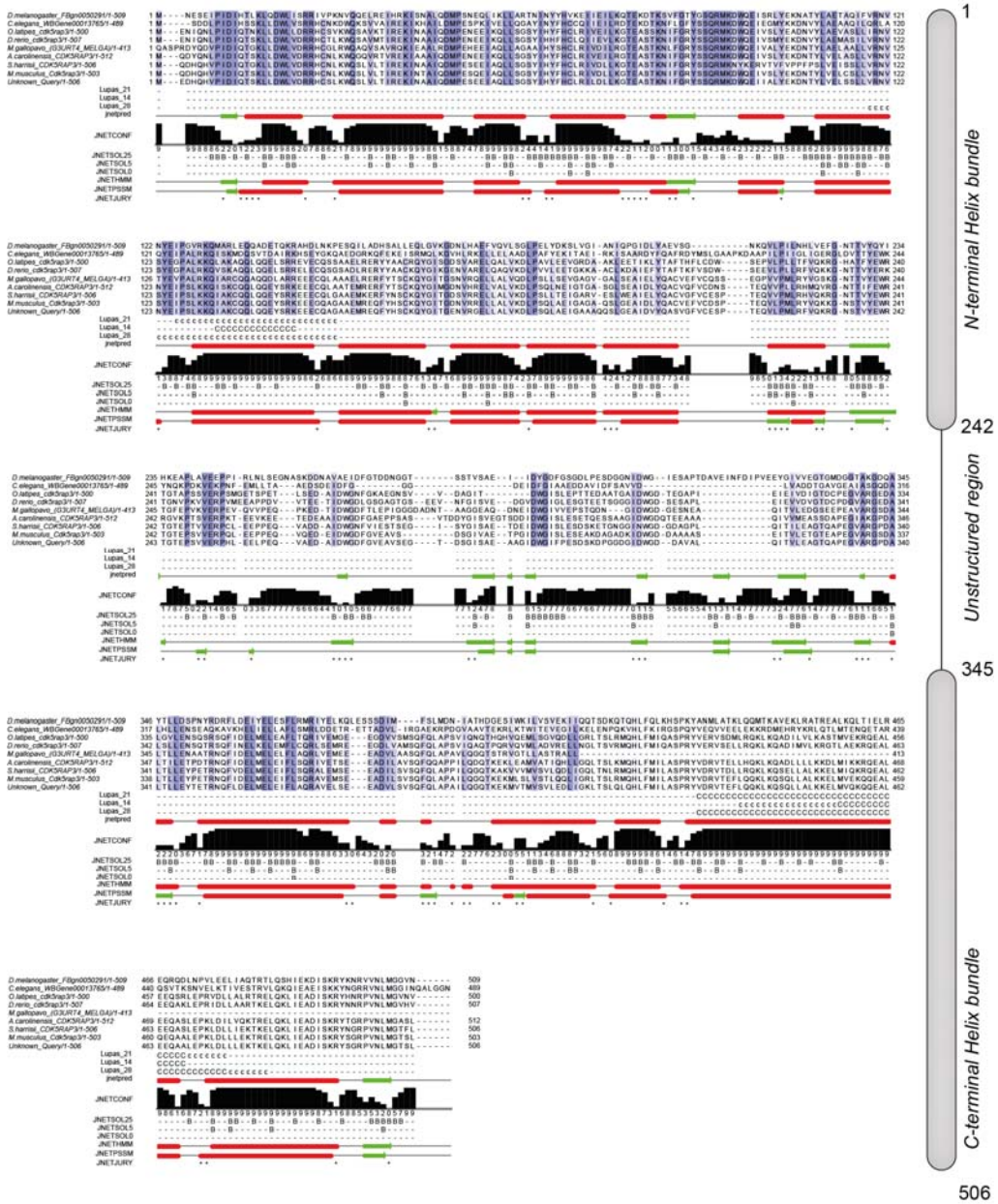



**Figure S6: - Sequence and secondary structure analysis of CDK5RAP3**

**A. (Left)** Multiple sequence alignment of CDK5RAP3 homologs showing predicted secondary structure features using JPred$^{TM}$ secondary structure program module from Jalview 2.11.1.7.**(Right)** Schematic showing predicted secondary structure features of CDK5RAP3.

## Figure S7

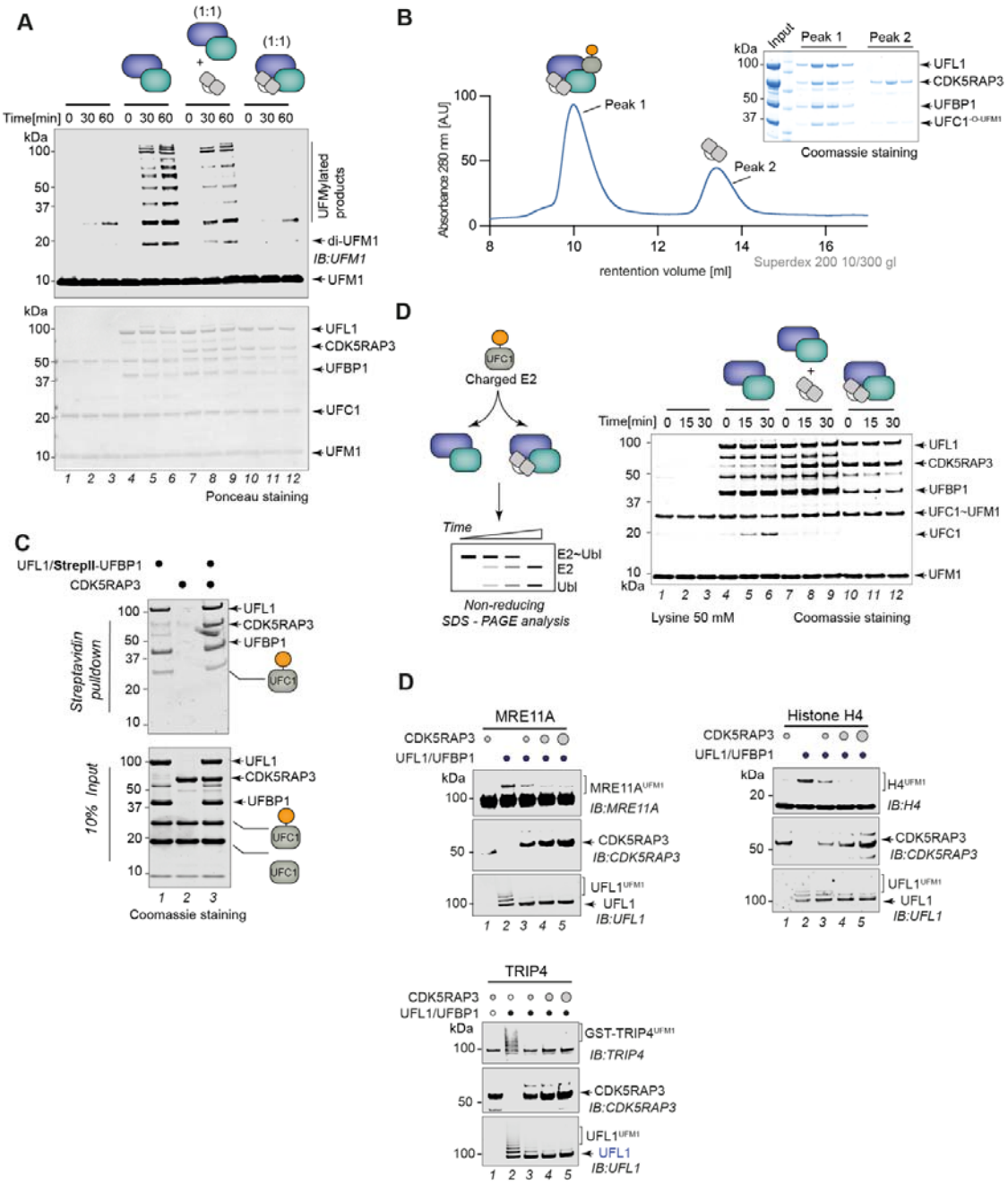



**Figure S7:- Role of CDK5RAP3 in UFMylation**

**A.** *In vitro* UFMylation assay to compare the E3 ligase activity of UFL1/UFBP1 mixed with CDK5RAP3 and preassembled ternary E3 ligase complex containing UFL1/UFBP1/CDK5RAP3.

**B.** UV traces of gel filtration chromatogram showing co-migration of (UFL1/UFBP1)/CDK5RAP3/UFC1$^{-O\text{-}UFM1}$ complex. Approximately, 200 µl of

sample containing (UFL1/UFBP1)/CDK5RAP3/UFC1$^{-O-UFM1}$ in the molar ratio of 1:1.5:3 was mixed and incubated at 4°C for 1 h and loaded onto a Superdex 200 10/300 g/l column. The fractions were collected and analysed on a 4-12% SDS PAGE and visualized by Coomassie staining.

**C.** Pulldown assay to check for interaction of UFL1/UFBP1 with charged UFC1 in the presence of absence of CDK5RAP3. Around 10 µM of Untagged UFC1 and 10 µM OF UFC1-O-UFM1 were mixed with 5 µM of UFL1/UFBP1 complex in the presence and absence of CDK5RAP3.

**D.** Lysine discharge assays to check for UFC1 discharge in the presence of UFL1/UFBP1 mixed with CDK5RAP3 and preassembled UFL1/UFBP1/CDK5RP3 complex. The reaction products were run on a 4-12% SDS PAGE analysis and visualized by LiCor.

**E.** *In vitro* UFMylation assay to monitor UFMylation of purified substrates namely TRIP4, MRE11A and Histone H4.

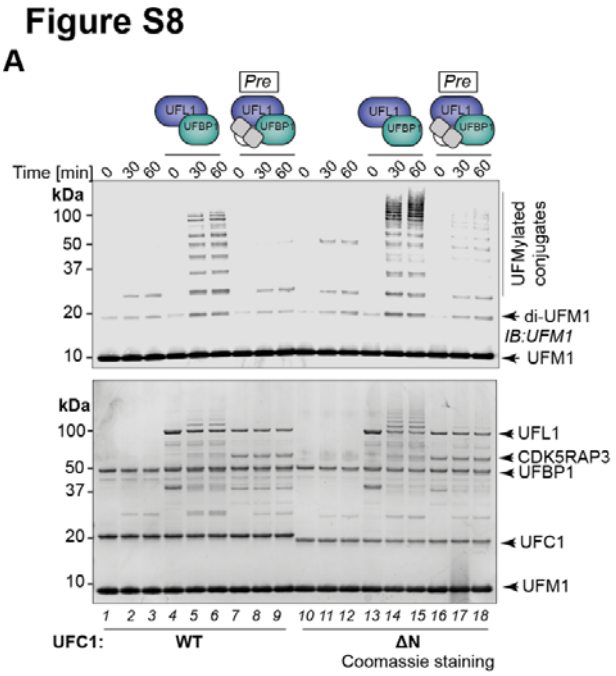



**Figure S8:- Role of the N-terminus of UFC1 in regulating UFMylation**

**A.** *In vitro* UFMylation assay to compare the activites of $UFC1^{WT}$ and $UFC1^{\Delta N}$ on its own, in the presence of UFL1/UFBP1 or in the presence of preassembled UFL1/UFBP1/CDK5RAP3 complex (Representative of three independent experiments).