Structural and Functional Analysis of a Novel Interaction Motif within UFM1-activating Enzyme 5 (UBA5) Required for Binding to Ubiquitin-like Proteins and Ufmylation*

Sabrina Habisov[‡§¶1], **Jessica Huber**[‖1], **Yoshinobu Ichimura**[**1], **Masato Akutsu**[§], **Natalia Rogova**[‖], **Frank Loehr**[‖], **David G. McEwan**[§‡‡], **Terje Johansen**[§§2], **Ivan Dikic**[§¶], **Volker Doetsch**[‖], **Masaaki Komatsu**[**3], **Vladimir V. Rogov**[‖4], and **Vladimir Kirkin**[‡5]

From ‡Translational Innovation Platform Oncology, Merck KGaA, Frankfurter Strasse 250, 64293 Darmstadt, Germany, the §Institute of Biochemistry II, Goethe University School of Medicine, Theodor-Stern-Kai 7, 60590 Frankfurt am Main, Germany, the ¶BMLS Buchmann Institute for Molecular Life Sciences, Goethe University Frankfurt, Riedberg Campus, Max-von-Laue-Strasse 15, 60438 Frankfurt am Main, Germany, the ‖Institute of Biophysical Chemistry and Center for Biomolecular Magnetic Resonance, Goethe University, Max-von-Laue-Strasse 9, 60438 Frankfurt am Main, Germany, the **Department of Biochemistry, School of Medicine, Niigata University, Chuo-ku, Niigata 951-8510, Japan, the ‡‡Division of Cell Signaling and Immunology, School of Life Sciences, University of Dundee, Dundee, DD1 4HN United Kingdom, and the §§Molecular Cancer Research Group, Institute of Medical Biology, University of Tromsø-The Arctic University of Norway, 9037 Tromsø, Norway

The covalent conjugation of ubiquitin-fold modifier 1 (UFM1) to proteins generates a signal that regulates transcription, response to cell stress, and differentiation. Ufmylation is initiated by ubiquitin-like modifier activating enzyme 5 (UBA5), which activates and transfers UFM1 to ubiquitin-fold modifier-conjugating enzyme 1 (UFC1). The details of the interaction between UFM1 and UBA5 required for UFM1 activation and its downstream transfer are however unclear. In this study, we described and characterized a combined linear LC3-interacting region/UFM1-interacting motif (LIR/UFIM) within the C terminus of UBA5. This single motif ensures that UBA5 binds both UFM1 and light chain 3/γ-aminobutyric acid receptor-associated proteins (LC3/GABARAP), two ubiquitin (Ub)-like proteins. We demonstrated that LIR/UFIM is required for the full biological activity of UBA5 and for the effective transfer of UFM1 onto UFC1 and a downstream protein substrate both *in vitro* and in cells. Taken together, our study provides important structural and functional insights into the interaction between UBA5 and Ub-like modifiers, improving the understanding of the biology of the ufmylation pathway.

UFM1[6] is a ubiquitin (Ub)-like protein (UBL) that shares structural, but little sequence, similarity with Ub and, like Ub, can be conjugated to lysine residues of itself or target proteins (1–3). UFM1 is activated by the E1 enzyme UBA5, which forms a high-energy thioester bond between its catalytic cysteine and the exposed C-terminal glycine of UFM1. Activated UFM1 is transferred to the catalytic cysteine of the E2 enzyme UFC1. Finally, a specific E3 ligase, UFL1 (UFM1 ligase 1, also known as RCAD), transfers UFM1 onto protein substrates (4–7). The two UFM1-specific proteases, UFSP1 and UFSP2, release UFM1 from its conjugated substrates and are additionally responsible for the post-translational processing of UFM1 to expose its C-terminal glycine (2).

Study of ufmylation (covalent conjugation of proteins with mono- or polyUFM1) is still in its infancy, but new details are beginning to emerge. For example, by covalently modifying the transcriptional co-activator ASC1, polyUFM1 chains enhance the association of ASC1 with p300 and SRC1 (steroid receptor coactivator 1), leading to the transcriptional activation of estrogen receptor α target genes and improved anchorage-independent growth of tumor cells (3). The role of UFBP1 (UFM1-binding protein 1 containing a PCI domain, also known as C20orf116 and DDRGK1) modification by UFM1 is less clear. However, it localizes to the endoplasmic reticulum (ER) and is involved in ER stress response (4, 7, 8). Interestingly, conditions that induce ER stress lead to increased expression of components of the UFM1 conjugation system, whereas their downregulation exacerbates ER stress and sensitizes cells to apoptosis (5–9). Of note, the ufmylation pathway is indispensable for erythro- and megakaryopoiesis, with $Uba5^{-/-}$, $Ufl1^{-/-}$, and

LIR/UFIM Is Required for UBA5 Function

$Ufbp1^{-/-}$ mice dying *in utero* because of severe anemia connected to the reduced numbers of functionally differentiated erythrocytes (5, 8, 10).

Ub and UBLs interact with specialized domains or short linear motifs present in the versatile group of proteins that effectively act as Ub and UBL receptors (11, 12). Through noncovalent interactions, receptor molecules are recruited to the Ub/UBL-conjugated proteins, regulating their localization and/or stability or mediating the assembly of molecular platforms responsible for downstream signal propagation. For example, the ubiquitin-associated (UBA) domain of the autophagic receptor p62/SQSTM1 recognizes ubiquitylated misfolded proteins and is required for their packing into inclusion bodies and subsequent lysosomal degradation (13, 14). In addition to its interaction with polyUb chains, p62/SQSTM1 interacts directly with the UBL proteins LC3/GABARAPs, which are conjugated to phosphatidylethanolamine enriched in autophagic membranes. Membrane-conjugated LC3/GABARAPs mediate lipid bilayer tethering and hemifusion (15, 16), drive expansion of autophagosomes, and via autophagic receptors, target autophagy cargo to the endolysosomal compartment (17). p62/SQSTM1 binds LC3/GABARAPs via a short linear sequence, designated the LC3-interacting region (LIR) and broadly defined by the core sequence (W/F/Y)XX(L/I/V), where X is any amino acid (aa) (13, 18). LIRs are increasingly found in proteins involved in signal transduction, such as ULK1-ATG13-FIP200 and KBTBD6/KBTBD7, as well as proteins in which turnover is regulated by autophagy (18–20).

Although UBA5 was originally identified as GABARAPL2/GATE-16-interacting protein (1), the molecular mechanisms behind this interaction with UBLs are not known. In the current study, we have identified a short linear motif in the C terminus of UBA5 that drives the interaction with either UFM1 or LC3/GABARAP proteins, defining this as a LIR/UFIM (for LC3-interacting region/UFM1-interacting motif). We present structural details for its interaction with UFM1 and model its interaction with GABARAPL2 and LC3B. By combining biophysical, biochemical and cellular techniques, we have provided a detailed characterization of the new UBL-binding motif and generated evidence for its role in the ability of UBA5 to mediate UFM1 conjugation *in vitro* and in cells. Our data suggest that a single UBL-binding element within UBA5 is responsible for both its function as an E1 enzyme in the ufmylation pathway and the interaction with LC3/GABARAP proteins.

Experimental Procedures

DNA Construction and Site-directed Mutagenesis—Plasmid generation was performed using either standard cloning procedures or in a Gateway cloning system (Thermo Fisher Scientific). Site-directed mutagenesis was performed using *Pfu*Ultra II fusion HS DNA polymerase (Agilent Technologies) according to the manufacturer's instructions. A comprehensive list of DNA constructs used in this study is found in Table 1.

Cell Culture and Transfection—HEK293 and U2OS cells were maintained in Dulbecco's modified Eagle's medium (Gibco) supplemented with 10% fetal bovine serum (Biochrome). Transient transfection of HEK293 and U2OS cells was performed for 24 h using X-tremeGENE 9 DNA transfec-

tion reagent (Roche) according to the manufacturer's instructions. *UBA5* knock-out HEK293 cells were generated by a CRISPR/Cas9 technology. UBA5 guide RNA was designed using CRISPR Design website, and it was subcloned into pX330-U6-chimeric_BB-CBh-hSpCas9 (Addgene 42230), a human codon-optimized SpCas9 and chimeric guide RNA expression plasmid. HEK293 cells were transfected with the vector together with pEGFP vector (pEGFP-C1; Clontech 6084-1) and cultured for 2 days. Thereafter, the EGFP-positive cells were sorted and expanded. Loss of *UBA5* was confirmed by heteroduplex mobility assay followed by immunoblot analysis with anti-UBA5 antibody.

Antibodies, Chemicals, and Peptides—The antibodies for Western blot and immunofluorescence staining used in this study are listed in Table 2. A stock solution of bafilomycin A1 (Calbiochem) was prepared in dimethyl sulfoxide (Sigma-Aldrich). The other reagents used are indicated in the text where necessary. Peptides used in this study were obtained from commercial sources and were stored on PVDF membranes or reconstituted in PBS or dimethyl sulfoxide. A list of peptides used in this study is found in Table 3.

Preparation of Peptides and Proteins for Isothermal Titration Calorimetry (ITC), NMR, and X-ray Studies—Human LC3/GABARAPs cloned under a modified Ub tag (21) were expressed and purified as described previously (20, 22, 23). UFM1 was cloned under a cleavable Ub19 leader (21) and expressed and purified in a similar manner. Short artificial tetrapeptides (GAMG- and GAME-) remained N-terminally attached to each purified protein after tobacco etch virus (TEV) cleavage. All proteins were equilibrated against a buffer (50 mM sodium phosphate, 100 mM NaCl, pH 7.0) prior to NMR and ITC experiments. For ITC and NMR titration experiments, hUBA5 LIR/UFIM peptide (EIIHEDNEWGIELVSEVSE; obtained from GenScript) was dissolved in the corresponding buffer.

Isothermal Titration Calorimetry—All ITC experiments were performed at 25 °C using a VP-ITC microcalorimeter (MicroCal Inc.) and analyzed with ITC Origin 7.0 software (MicroCal Inc.) based on a "one-site" binding reaction. For binding studies of hUBA5 LIR/UFIM peptide to UBLs, 500 μM peptide was titrated to 20–25 μM UBL. ITC titrations of hUBA5 LIR/UFIM and p62/SQSTM1 LIR peptides into UFM1 were performed in the same way (500 μM UBA5 LIR/UFIM and p62/SQSTM1 LIR peptides were titrated to 25 μM UFM1).

Nuclear Magnetic Resonance—All NMR experiments were performed at 25 °C on Bruker Avance spectrometers operating at proton Larmor frequencies of 500, 700, 800, and 900 MHz and were analyzed using Sparky 3.114 software (University of California, San Francisco). Backbone ^{1}H and ^{15}N resonances and side-chain ^{13}C resonances of GABARAPL2 and UFM1 were assigned using a [^{15}N-^{1}H]BEST-TROSY version of three-dimensional HNCACB (24) and a [^{15}N-^{1}H]TROSY version of (H)C(CCO)NH-TOCSY (25, 26) experiments with 0.8–1.0 mM uniformly ^{13}C,^{15}N-labeled protein samples. For NMR titration experiments, the nonlabeled hUBA5 LIR/UFIM peptide was titrated to 150 μM ^{15}N-LC3B, 160 μM ^{15}N-GABARAPL2, or 180 μM ^{15}N-UFM1 to a final molar ratio of 1:1.5 for GABARAPL2 or 1:4 in the case of LC3B and UFM1. For the NMR competition

TABLE 1
Plasmids used in this study

Amp, ampicillin; Bla, blasticidin; Cm, chloramphenicol; fwd, forward; h, human; Hygro, hygromycin; Kan, kanamycin; mut, mutagenesis; Neo, neomycin; Puro, puromycin; Fwd, forward; Rev, reverse; NA, not applicable.

Construct	Utility	Source/Reference	Resistance marker	Restriction sites	Oligonucleotides used for cloning
pDestEGFP-GABARAPL2	Mammalian expression of EGFP-GABARAPL2	Ref. 13	Neo	NA	NA
pDestEGFP-LC3B	Mammalian expression of EGFP-LC3	Ref. 13	Neo	NA	NA
pcDNA3-Myc-UFBP1	Mammalian expression of Myc-hUFBP1	Ref. 1	Amp/Neo	BamHI/XhoI	NA
pcDNA3-Myc-UFL1	Mammalian expression of Myc-hUFL1	Ref. 1	Amp/Neo	BamHI/XhoI	NA
pcDNA3-Myc-UFM1ΔC2	Mammalian expression of Myc-hUFM1ΔC2	Ref. 1	Amp/Neo	BamHI/XhoI	NA
pcDNA5.1-HA-GW	Destination vector for Gateway cloning system; expression of N-terminal HA fusions in mammalian cells	This study	Amp/Hygro	NA	NA
pcDNA5.1-HA-GW-UBA5	Mammalian expression of human HA-UBA5	This study	Amp/Hygro	NA	NA
pcDNA5.1-HA-GW-UBA5 (1–330 aa)	Mammalian expression of human HA-UBA5 (1–330 aa) mutant	This study	Amp/Hygro	NA	NA
pcDNA5.1-HA-GW-UBA5 (ΔLIR/UFIM)	Mammalian expression of human HA-UBA5 (ΔLIR/UFIM) mutant	This study	Amp/Hygro	NA	NA
pcDNA5.1-HA-GW-UBA5 (W341A/L345A)	Mammalian expression of human HA-UBA5 (W341A/L345A) mutant	This study	Amp/Hygro	NA	NA
pcDNA5.1-HA-GW-ATG4D	Mammalian expression of human HA-ATG4D mutant (negative control in this study)	This study	Amp/Hygro	NA	NA
pDEST15	Destination vector for Gateway cloning system; expression of N-terminal GST fusions in bacteria	Thermo Fisher	Cm	NA	NA
pDEST15-hUBA5	Bacterial expression of GST-hUBA5	This study	Cm	Gateway cloning	NA
pDEST15-hUBA5 (W341A/L345A)	Bacterial expression of GST-hUBA5 (W341A/L345A)	This study	Cm	Gateway cloning	NA
pDEST15-hUBA5 (1–330 aa)	Bacterial expression of GST-hUBA5 (1–330 aa)	This study	Cm	Gateway cloning	NA
pDEST15-hUBA5 (ΔLIR/UFIM)	Bacterial expression of GST-hUBA5 (ΔLIR/UFIM = Δ340–347)	This study	Cm	Gateway cloning	NA
pENTR1A	Entry vector for Gateway cloning system	Thermo Fisher	Kan	NA	NA
pENTR1A-hUFM1	Cloning human UFM1 in Gateway system	This study	Kan	NotI/EcoRI	Fwd, CGCGGCCGCTCGAAGGTTTCCTTTAAGATCACG; Rev, CGATATCTTAACAACTTCCAACACGATCT
pENTR1A-hUBA5	Cloning human UBA5 in Gateway system	This study	Kan	NotI/EcoRI	Fwd, CGCGGCCGCGGAGTCTGTGGAGCGCC; Rev, CGATATCTTCTACATATTCTTCATTTTGGCCATG
pENTR1A-hUBA5 (W341A/L345A)	Cloning human UBA5 (W341A/L345A) in Gateway system	This study	Kan	NotI/EcoRI	Fwd (mut), CTTCTGAAACCTCAGATACCGCCTCAATACCC GCTTCAATTATCTTCATGGATTATCTC; Rev (mut), GAGATAATCCATGAAGATAATGAAGCGGGTAT TGAGGCGGTATCTGAGGTTTCAGAAG
pENTR1A-hUBA5 (1–330 aa)	Cloning human UBA5 (1–330 aa) in Gateway system	This study	Kan	NotI/EcoRI	Fwd, CGCGGCCGCGCGGAGTCGTGGAGCGCC; Rev, CGCGGCCGCGAGGAAGAGATAATCCATGAAGATA
pENTR1A-hUBA5 (ΔLIR/UFIM)	Cloning human UBA5 (ΔLIR/UFIM = Δ340–347) in Gateway system	This study	Kan	NotI/EcoRI	Fwd (mut), GAGGAAGAGATAATCCATGAAGATAATGAG GTTTCAGAAGAGGAACTGAAAAATTTTTCAG; Rev (mut), CTGAAAAATTTTTCAGTTCCTCTTCTGAAACC TCATTATCTTCATGGATTATCTCTTCCTC
pETM60-Ub3	Bacterial expression of proteins with a modified ubiquitin tag	Ref. 20	Kan	NcoI/BamHI	NA
pETM60-Ub3-hLC3A	Bacterial expression of LC3A with a modified ubiquitin tag	Ref. 20	Kan	NcoI/BamHI	NA
pETM60-Ub3-hLC3B, -hGABARAPL1	Bacterial expression of LC3B with a modified ubiquitin tag	Ref. 20	Kan	NcoI/BamHI	NA
pETM60-Ub3-hGABARAPL1	Bacterial expression of GABARAPL1 with a modified ubiquitin tag	Ref. 20	Kan	NcoI/BamHI	NA
pET39b-Ub19	Bacterial expression of proteins with a modified ubiquitin tag	Ref. 20	Kan	NcoI/BamHI	NA
pET39b-Ub19-LC3C	Bacterial expression of proteins with a modified ubiquitin tag	Ref. 20	Kan	NcoI/BamHI	NA
pET39b-Ub19--hGABARAP	Bacterial expression of proteins with a modified ubiquitin tag	20	Kan	NcoI/BamHI	NA
pET39b-Ub19-hGABARAPL2	Bacterial expression of proteins with a modified ubiquitin tag	Ref. 20	Kan	NcoI/BamHI	NA
pET39b-Ub19-hUFM1(2–83)	Bacterial expression of human UFM1 (2–83) with a modified ubiquitin tag	This study	Kan	NcoI/BamHI	Fwd, CCATGGAATCGAAGGTTTCCTTTAAGATCACG; Rev, GATTATTCCTAGAGATCGTGTTGGATGAGGATCC
pGEX-4T-1	Bacterial expression of N-terminally tagged GST fusion proteins	GE Biosciences	Amp	NA	NA
pGEX-4T-1-hGABARAP(ΔG)	Bacterial expression of GST-hGABARAP (ΔG)	Ref. 13	Amp	BamHI/XhoI	NA
pGEX-4T-1-hGABARAPL1(ΔG)	Bacterial expression of GST-hGABARAPL1 (ΔG)	Ref. 13	Amp	EcoRI/XhoI	NA
pGEX-4T-1-hGABARAPL2(ΔG)	Bacterial expression of GST-hGABARAPL2 (ΔG)	Ref. 13	Amp	EcoRI/XhoI	NA
pGEX-4T-1-hMAP1LC3A(ΔG)	Bacterial expression of GST-hLC3A (ΔG)	Ref. 13	Amp	EcoRI/XhoI	NA
pGEX-4T-1-hMAP1LC3B(ΔG)	Bacterial expression of GST-hLC3B (ΔG)	Ref. 13	Amp	EcoRI/XhoI	NA
pGEX-4T-1-hMAP1LC3C(ΔG)	Bacterial expression of GST-hLC3C (ΔG)	Ref. 13	Amp	EcoRI/XhoI	NA
pGEX-4T-1-hUFM1(ΔG)	Bacterial expression of GST-hUFM1 (ΔG)	Ref. 13	Amp	BamHI/NotI	NA
pGEX6p-1	Bacterial expression of N-terminally tagged GST fusion proteins	GE Healthcare Biosciences	Amp	NA	NA
pGEX6p-1-hUBA5(W341A)	Bacterial expression of GST-hUBA5 (W341A)	This study	Amp	BamHI/XhoI	Fwd, CGGGATCCATGGCCGAGTCTGTGGAGCG; Rev, CCGCTCGAGCTACATATTCTTCATTTTGG

TABLE 1—continued

Construct	Utility	Source/Reference	Resistance marker	Restriction sites	Oligonucleotides used for cloning
pGEX6p-1-hUBA5(G342A)	Bacterial expression of GST-hUBA5 (G342A)	This study	Amp	BamHI/XhoI	Fwd, CGGGATCCATGGCGGAGTCTGTGGAGCG; Rev, CCGCTCGAGCTACATATTCTTCATTTTGG
pGEX6p-1-hUBA5(W341A/L345A)	Bacterial expression of GST-hUBA5 (W341A/L345A)	This study	Amp	BamHI/XhoI	Fwd, CGGGATCCATGGCGGAGTCTGTGGAGCG; Rev, CCGCTCGAGCTACATATTCTTCATTTTGG
pGEX6p-1-hUBA5(ΔLIR/UFIM)	Bacterial expression of GST-hUBA5 (ΔLIR/UFIM)	This study	Amp	BamHI/XhoI	Fwd, CGGGATCCATGGCGGAGTCTGTGGAGCG; Rev, CCGCTCGAGCTACATATTCTTCATTTTGG
pGEX6p-1-hUBA5(1–330 aa)	Bacterial expression of GST-hUBA5 (1–330 aa)	This study	Amp	BamHI/XhoI	Fwd, CGGGATCCATGGCGGAGTCTGTGGAGCG; Rev, CCGCTCGAGCTATTCTTGTATAACCTCTTG
pGEX6p-1-hLC3B	Bacterial expression of GST-hLC3B (G120 end)	Ref. 36	Amp	BamHI/XhoI	NA
pGEX6p-2-hUBA5	Bacterial expression of GST-hUBA5	Ref. 1	Amp	BamHI/XhoI	NA
pGEX6p-2-hUFC1	Bacterial expression of GST-hUFC1	Ref. 1	Amp	BamHI/XhoI	NA
pGEX6p-2-hUFM1ΔC2	Bacterial expression of GST-hUFM1ΔC2	Ref. 1	Amp	NA	NA
pIRES-3xFLAG	Mammalian expression of FLAG-tagged proteins	This study	Amp/Neo	BamHI/XhoI	NA
pIRES-3xFLAG-hUBA5	Mammalian expression of FLAG-hUBA5	This study	Amp/Neo	BamHI/XhoI	Fwd, CGGGATCCATGGCGGAGTCTGTGGAGCG; Rev, CCGCTCGAGCTACATATTCTTCATTTTGG
pIRES-3xFLAG-hUBA5(W341A)	Mammalian expression of FLAG-hUBA5 (W341A)	This study	Amp/Neo	BamHI/XhoI	Fwd, CGGGATCCATGGCGGAGTCTGTGGAGCG; Rev, CCGCTCGAGCTACATATTCTTCATTTTGG
pIRES-3xFLAG-hUBA5(G342A)	Mammalian expression of FLAG-hUBA5 (G342A)	This study	Amp/Neo	BamHI/XhoI	Fwd, CGGGATCCATGGCGGAGTCTGTGGAGCG; Rev, CCGCTCGAGCTACATATTCTTCATTTTGG
pIRES-3xFLAG-hUBA5(W341A/L345A)	Mammalian expression of FLAG-hUBA5 (W341A/L345A)	This study	Amp/Neo	BamHI/XhoI	Fwd, CGGGATCCATGGCGGAGTCTGTGGAGCG; Rev, CCGCTCGAGCTACATATTCTTCATTTTGG
pIRES-3xFLAG-hUBA5(ΔLIR/UFIM)	Mammalian expression of FLAG-hUBA5 (ΔLIR/UFIM)	This study	Amp/Neo	BamHI/XhoI	Fwd, CGGGATCCATGGCGGAGTCTGTGGAGCG; Rev, CCGCTCGAGCTATTCTTGTATAACCTCTTG
pIRES-3xFLAG-hUBA5(1–330 aa)	Mammalian expression of FLAG-hUBA5 (1–330 aa)	This study	Amp/Neo	BamHI/XhoI	Fwd, CGGGATCCATGGCGGAGTCTGTGGAGCG; Rev, CCGCTCGAGCTACATATTCTTCATTTTGG
pmCherry-C1	Mammalian expression of proteins fused to the C-terminus of mCherry	Clontech	Kan/Neo	NA	NA
pmCherry-GABARAP	Mammalian expression of mCherry-GABARAP	This study	Kan/Neo	XhoI-BamHI	NA

assay, a ^{15}N-labeled UFM1 sample (200 μM) was supplied with a double molar excess of the nonlabeled hUBA5 LIR/UFIM peptide (400 μM) and incubated at 25 °C. Nonlabeled GABARAPL2 was added to this mixture to a final concentration of 400 μM to have a 1:1 ratio to the hUBA5 LIR/UFIM and 800 μM to have a 1:2 ratio to the hUBA5 LIR/UFIM.

X-ray Crystallography—Purified UFM1 was dialyzed into a crystallization buffer (50 mM Tris, pH 8.0, 100 mM NaCl). The synthetic hUBA5 LIR/UFIM peptide was dissolved directly in the same buffer. Crystals of UFM1 in complex with the hUBA5 LIR/UFIM peptide were obtained using 1.6 M sodium phosphate and 0.1 M HEPES, pH 7.5, as a reservoir solution by the sitting-drop vapor diffusion method at 293 K. Diffraction data were collected at the Swiss Light Source, beamline PXIII, and processed with XDS (27). The crystal structure of the complex was determined by molecular replacement using the UFM1 NMR structure (PDB ID: 1J0G) as a search model. Manual model building and refinement were done with Coot, the CCP4 software suite, and Phenix (28–30). The final statistics of refined models are shown in Table 4.

Preparation of Proteins for Interaction Assays—GST fusion proteins were expressed from pGEX-4T1 vectors in BL21-competent *Escherichia coli* (Thermo Fisher Scientific) in LB medium (VWR) at 18 °C overnight with 1 mM isopropyl 1-thio-β-D-galactopyranoside (EMD Millipore). Bacteria were harvested, and pellets were frozen at −80 °C for 24 h. Bacterial pellets were lysed after resuspension in a lysis buffer (50 mM Tris, 100 mM NaCl, 5% glycerol, 0.2 mg/ml chicken egg lysozyme, 1 mM PMSF, and 1 μg/ml DNase) by sonication (TT13 Sonotrode, 25% amplitude, for 12 min with a 0.5/0.5-s pulse). Lysates were centrifuged at $15,000 \times g$ at 4 °C for 30 min. The resulting supernatants were incubated with glutathione beads (GE Healthcare, 1 ml of 50% slurry/1 liter of *E. coli* culture) with rotation for 1 h at 4 °C. The beads were washed three times with lysis buffer. Bound GST fusion proteins were eluted twice with 10 mM glutathione, each time for 20 min, at 4 °C. Elution buffer was then exchanged to 50 mM Tris, 100 mM NaCl by dialysis, and proteins were concentrated using Centricon concentrators (GE Healthcare). For the competitive GST pull-down assay, UFM1 and GABARAPL2 were expressed from pET39-Ub19 vectors in T7 Express-competent *E. coli* (New England Biolabs) in LB medium at 18 °C overnight after induction with 1 mM isopropyl 1-thio-β-D-galactopyranoside. The bacterial culture was processed as described above. Purified proteins were dialyzed overnight, and Ub tags were removed by tobacco etch virus cleavage. The cleaved tags were removed by binding to Talon beads (GE Healthcare) at 4 °C with rotation for 1 h. The supernatant was concentrated using Centricon concentrators.

GST Overlay Assay Using Peptide Arrays—The peptide array scan of full-length UBA5 was performed exactly as described previously (19). PVDF membranes containing arrays of spotted synthetic peptides (5 nM, made by JPT) were activated for 10 s in 100% methanol and then rinsed under running tap water followed by blocking in 5% skim milk in TBS-T (50 mM Tris, 150 mM NaCl, 0.2% Tween, pH 7.6) for 1 h at room temperature with shaking. Purified GST fusion proteins were diluted in TBS-T to a final concentration of 1 μg/ml. Membranes were

TABLE 2

Antibodies used in this study

IF, immunofluorescence; WB, Western blotting; NA, not applicable.

Antibody	Supplier	Catalog No.	Host	Clonality	Application (dilution)
Primary					
Anti-FLAG (M2)	Sigma-Aldrich	F1804	Mouse	Monoclonal	WB (1:1000)
Anti-GABARAPL2	Abcam	ab126607	Rabbit	Monoclonal	WB (1:1000)
Anti-GFP	Roche/Sigma-Aldrich	11814460001	Mouse	Monoclonal	WB (1:200)
Anti-GST (HRP)	GE Healthcare	RPN1236	NA	NA	WB (1:5000)
Anti-HA	Santa Cruz Biotechnology	sc-7392	Mouse	Monoclonal	WB (1:1000)
Anti-HA (3F10)	Roche	112165071	Rat	Monoclonal	IF (1:500)
Anti-Ufm1	Abcam	ab109305	Rabbit	Monoclonal	WB (1:1000)
Secondary					
Goat anti-rat IgG-Cy3	Jackson ImmunoResearch	A11001	Goat	Polyclonal	IF (1:200)
ECL anti-mouse (HRP)	GE Healthcare	NA931V	Sheep	Polyclonal	WB (1:2000)
ECL anti-rabbit (HRP)	GE Healthcare	NA934	Donkey	Polyclonal	WB (1:5000)

TABLE 3

Peptides used in this study

Mutations introduced in comparison with the wild-type UBA5 LIR/UFIM peptide are underlined. All peptides were obtained from GeneScript. A GST overlay assay was used for all peptides except hUBA5 LIR/UFIM.

Peptide name	Sequence
1	QEVIQEEEEIIHEDNEWGIE
2	IQEEEEIIHEDNEWGIELVS
3	EEEIIHEDNEWGIELVSEVS
4	IIHEDNEWGIELVSEVSEEE
5	EDNEWGIELVSEVSEEELKN
6	EWGIELVSEVSEEELKNFSG
7	IELVSEVSEEELKNFSGPVP
8	VSEVSEEELKNFSGPVPDLP
9	VSEEELKNFSGPVPDLPEGI
10	EELKNFSGPVPDLPEGITVA
11	KNFSGPVPDLPEGITVAYTI
12	SGPVPDLPEGITVAYTIPKK
13	VPDLPEGITVAYTIPKKQED
14	LPEGITVAYTIPKKQEDSVT
15	GITVAYTIPKKQEDSVTELT
16	VAYTIPKKQEDSVTELTVED
17	TIPKKQEDSVTELTVEDSGE
18	KKQEDSVTELTVEDSGESLE
19	EDSVTELTVEDSGESLEDLM
20	VTELTVEDSGESLEDLMAKM
21	LTVEDSGESLEDLMAKMKNM
Peptide 1	AIHENDNEWGIELVSE
Peptide 2	IAHENDNEWGIELVSE
Peptide 3	IIAENDNEWGIELVSE
Peptide 4	IIHANDNEWGIELVSE
Peptide 5	IIHEADNEWGIELVSE
Peptide 6	IIHENANEWGIELVSE
Peptide 7	IIHENDAEWGIELVSE
Peptide 8	IIHENDNAWGIELVSE
Peptide 9	IIHENDNEAGIELVSE
Peptide 10	IIHENDNEWAIELVSE
Peptide 11	IIHENDNEWGAELVSE
Peptide 12	IIHENDNEWGIALVSE
Peptide 13	IIHENDNEWGIEAVSE
Peptide 14	IIHENDNEWGIELASE
Peptide 15	IIHENDNEWGIELVAE
Peptide 16	IIHENDNEWGIELVSA
Peptide 17	IIHENDNEWGIELVSE
Peptide 18	IIHENDNEFGIELVSE
Peptide 19	IIHENDNEYGIELVSE
Peptide 20	IIHENDNAEGIELVSE
Peptide 21	IIHENDNAFGIELVSE
Peptide 22	IIHENDNAYGIELVSE
Peptide 23	IIHENDNAAGIELVSE
hUBA5 LIR/UFIM[a]	EIIHEDNEWGIELVSE

[a] Assayed by ITC, NMR, and x-ray crystallography.

incubated for 2 h at room temperature with shaking. After four washes with TBS-T, membranes were incubated overnight at 4 °C with anti-GST antibody diluted (1:5000) in the blocking buffer. Membranes were washed three times for 10 min each in TBS-T followed by two washes with PBS. Antibodies bound to GST fusion protein-peptide complexes were detected using ECL detection reagent (GE-Healthcare) and a Bio-Rad Imager.

GST Pulldown Assay—Purified GST fusion proteins were incubated with glutathione-Sepharose 4B (GE Healthcare) for 1 h at 4 °C with rotation. The beads were washed three times in wash buffer (50 mM Tris, 100 mM NaCl, 5% glycerol). HEK293 cells were transiently transfected for 24 h with X-tremeGENE 9 transfection reagent (Roche) according to the manufacturer's instructions. Cells were lysed in a lysis buffer (150 mM NaCl, 5 mM EDTA, 1% Nonidet P40, 0.5% Na-deoxycholate, 50 mM Tris, pH 7.8, 0.1% SDS) and centrifuged at $20,000 \times g$, and the supernatant was transferred to fresh Eppendorf tubes. The protein concentration was determined using a Bradford assay (Bio-Rad). 200 μg of total cell lysate was incubated with beads for 1 h at 4 °C with rotation. The beads were washed three times in wash buffer, and proteins were eluted with the $1\times$ LDS loading dye (Thermo Fisher Scientific) containing 5 mM β-mercaptoethanol (Bio-Rad). For the *in vitro* competition assay, 2 μM GST-UBA5 immobilized on glutathione beads was incubated with 100 μM UFM1 for 1 h at 4 °C with rotation to achieve saturation. Samples were washed three times in wash buffer and incubated with varying amounts of GABARAPL2 (0.01 nM–100 μM) for 1 h at 4 °C with rotation. After three washings, immunoblot analysis was performed.

Co-immunoprecipitation Assay—HA/GFP fusion proteins were transiently expressed in HEK239 cells. Cell lysates were prepared after 24 h of expression using a modified radioimmune precipitation assay buffer (50 mM HEPES, pH 7.8, 150 mM NaCl, 1 mM EDTA, 1 mM EGTA, 10% glycerol, 1% Triton X-100, 25 mM NaF, 10 μM ZnCl$_2$). The protein concentration was determined by the Bradford assay, and 200 μg of whole lysate was incubated with 15 μl of anti-HA affinity matrix (Roche 11815016001)/15 μl of GFP-Trap-agarose beads (Chromotech) for 1 h with rotation at 4 °C. Beads were washed three times and then loaded on SDS-PAGE for analysis.

Immunoblot Analysis—Protein samples were resolved on 4–12% BisTris gels (Thermo Fisher Scientific) with MES running buffer (50 mM MES, 50 mM Tris, 0.1% SDS, 1 mM EDTA, pH 7.3) for 40 min at 200 V. Proteins were transferred onto PVDF membranes (Bio-Rad) by wet electroblotting at 55 mA for 90 min. Membranes were blocked for 1 h at room temperature with 5% BSA in TBS-T buffer. Primary antibodies were diluted in 5% BSA in TBS-T buffer, and membranes were incu-

bated overnight at 4 °C with rotation. After three washes with TBS-T (50 mM Tris, 150 mM NaCl, 0.2% Tween, pH 7.6), the membranes were incubated with secondary antibodies for 1 h at room temperature with shaking. Image acquisition was performed using an ECL detection reagent (GE Healthcare) and a Bio-Rad Chemidoc imaging system.

In Vitro Thioester Formation Assay—GST-UFM1ΔC2 (C-terminal glycine exposed form), GST-UBA5, GST-UBA5(W341A), GST-UBA5(G342A), GST-UBA5 (W341A/L345A), GST-UBA5 (ΔLIR), GST-UBA5(1–330 aa), GST-UFC1, and GST-GABARAPL2 were expressed in BL21(DE3) *E. coli* and purified using glutathione-Sepharose 4B (GE Healthcare). Recombinant proteins were eluted by cleaving GST with PreScission protease (GE Healthcare) on the column and dialyzed against 50 mM BisTris, pH 6.5, 100 mM NaCl, 10 mM MgCl$_2$, and 0.1 mM DTT (reaction buffer). The UFM1-UBA5 thioester formation assay was carried out in the reaction buffer containing 1.0–2.5 µg of UBA5 or UBA5 mutants, 0.5–1.0 µg of UFM1ΔC2, and 0.1 mM ATP. The UFM1-UFC1 thioester formation assay was assayed in the above reaction mixture supplemented with 1.0–3.0 µg of UFC1. The mixtures were incubated at 25 °C for 5 min (or other times as described in the figure legends), and then nonreducing NuPAGE LDS sample buffer (Thermo Fisher Scientific) was added to stop the reaction. The samples were subjected to NuPAGE followed by Coomassie Brilliant Blue staining. The signal intensities of the bands were quantified using NIH Image.

Statistical Analysis—Statistical analyses were performed using the unpaired *t* test (Welch test), and *p* values less than 0.05 were considered significant.

Results

UBA5 Interacts with UFM1 and LC3/GABARAP Proteins via an Evolutionary Conserved Combined LIR/UFIM Motif—We and others have reported previously that the UFM1-activating enzyme UBA5 interacts with GABARAPL2, also known as GATE-16 (1, 31), and with UFM1 (1, 32). However, it remained unclear which regions of UBA5 interacted with both UBLs. First, we performed multiple sequence alignment to determine the regions of UBA5 that are most highly conserved. Notably, three regions of high conservation were observed: the catalytic adenylation domain (found in the N-terminal part of the protein) and two poorly characterized regions (found at the C terminus of UBA5) (Fig. 1*A*). Although the C-terminal part of UBA5 was implicated in binding to UFM1 and UFC1 (32, 33), we noticed that the 14-aa-long sequence IHEDNEWGIELVSE (aa 335–348 in human UBA5) contains a stretch of amino acids remotely resembling the LIR consensus sequence (Fig. 1*B*). To elucidate this further in an unbiased approach, we turned to a peptide array analysis (19). We performed a peptide walk covering the complete sequence of human UBA5 using 20-mer peptides with three-amino acid intervals. The arrays were probed with recombinant GST or GST-GABARAP, and binding was detected with anti-GST antibodies. This way, a single LIR was identified in UBA5 with the 8-aa-long sequence EWGIELVS (Fig. 1*C*). This stretch contains the highly conserved Trp-341, Ile-343, Leu-345, and Val-346 residues, with the important feature that the spacing between the aromatic

and aliphatic residues does not conform to the canonical LIR (one intermediate residue instead of the usual two (Fig. 1*B*)).

We then tested whether all members of the LC3/GABARAP family of UBLs interact with UBA5 using GST-tagged LC3/GABARAPs and overexpressed human GFP-UBA5, the wild type, or its putative LIR mutants (Fig. 1*D*). Interestingly, UBA5 interacted preferentially with the GABARAP subfamily (Fig. 1, *E* and *F*) and weakly with LC3A (Fig. 1*E*). Recombinant UFM1, as well as UFC1, efficiently precipitated UBA5 (Fig. 1, *E* and *F*). Surprisingly, UBA5 LIR mutants showed significantly reduced levels of interaction with both LC3/GABARAPs and UFM1 (Fig. 1, *E* and *F*). Based on this double suppression effect, we designated the short region spanning aa 337 to 348 in human UBA5 as a combined LIR/UFIM. Indeed, deletion of LIR/UFIM (ΔLIR/UFIM), mutation of the key residues to alanine (W341A/L345A), or complete removal of the C terminus (1–330 aa) resulted in a loss of UBL interaction (Fig. 1, *E* and *F*). In contrast, UFC1 could still bind UBA5(ΔLIR/UFIM) and UBA5(W341A/L345A) mutants but not the one lacking the whole C terminus (1–330 aa) (Fig. 1*F*).

Given the potential biological significance of the observed UBA5-UBL interactions, we tested whether HA-tagged UBA5 would also interact with endogenous UBLs. Although UFM1 was readily co-precipitated from HEK293 cell lysates, no signal for endogenous GABARAPL2 was identified in samples containing proteins co-immunoprecipitated with HA-UBA5 (Fig. 1*G*). Treatment with bafilomycin A1, which induces the accumulation of autophagosomes, did not result in an improved interaction in cells (data not shown). This is in line with the co-localization experiments in which we observed no co-localization between UBA5 and GABARAP or GABARAPL2 in U2OS cells (data not shown). On the other hand, UBA5 LIR/UFIM mutants failed to precipitate endogenous UFM1 from HEK293 cell lysates (Fig. 1*G*), indicating the significance of LIR/UFIM for UBA5-UBL interactions within the cytoplasm. Thus, UBA5 interacts with both UFM1 and LC3/GABARAPs via a unique combined LIR/UFIM; however, under basal conditions in cells, UBA5 seems to preferentially interact with UFM1 but not with LC3/GABARAPs.

Characterization of the UBA5 LIR/UFIM-UBL Interaction—As truncation and mutation of LIR/UFIM completely abolished the interaction of UBA5 with both UFM1 and LC3/GABARAPs, we decided to characterize this sequence in detail. We again used an array of synthetic 20-mer peptides in a GST overlay assay (19) and probed for interaction with UFM1 and GABARAPL2. Surprisingly, both GST-UFM1 and GST-GABARAPL2 interacted efficiently only with the peptides containing the sequence EWGIELVS (Fig. 2*A*). Notably, binding to UFM1 shows a sequence dependence very similar to that of GABARAPL2, as demonstrated by the mutational (alanine substitution) analysis of LIR/UFIM peptides. Thus, alanine substitutions of the conserved Trp-341, Ile-343, and Leu-345 fully abolished peptide interaction with UFM1 (Fig. 2*B*). In the case of GABARAPL2, the inhibitory effects of the same mutations are obvious, albeit less pronounced because of the higher affinity of the interaction (see ITC data below). Mutations that affect both the UFM1 and GABARAPL2 interactions include Trp-341 and Val-346, with W341A having the great-
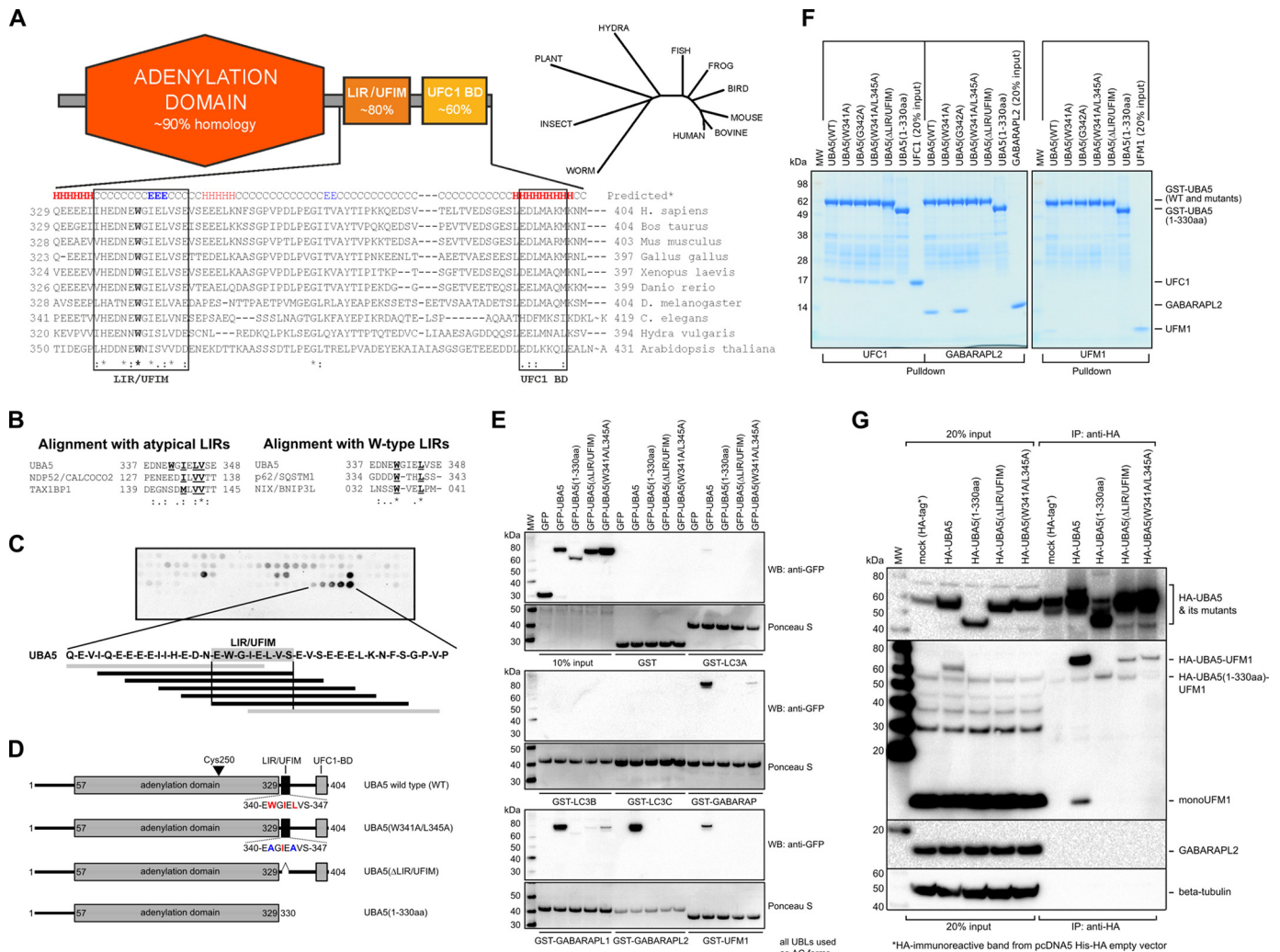
FIGURE 1. **UBA5 interacts with UFM1 and LC3/GABARAP proteins via an evolutionary conserved combined LIR/UFIM motif.** *A*, multiple sequence alignment of UBA5 C termini from different species (ClustalW2 algorithm). The degrees of homology (identity plus similarity) for domains are given as percentages (%) for each region. *Asterisk* (*), fully conserved residue; *colon* (:), strongly similar conservation; *period* (.), weakly similar conservation. The phylogenetic tree (*upper right plot*) was built for the indicated phylogenetic groups according to the "greatest likelihood" algorithm. UBA5 secondary structure elements were predicted by JPRED algorithms. *B*, alignment of LIR/UFIM sequences from human UBA5 with published LIRs from indicated human proteins. Multiple sequence alignment was performed using the ClustalW2 algorithm. *C*, identification of a GABARAP-interacting LIR motif in UBA5. Arrays of 20-mer peptides covering full-length UBA5 were synthesized and prepared on cellulose membranes. Each peptide was shifted three amino acids relative to the previous peptide. The arrays were probed with GST-GABARAP, and binding was detected with anti-GST antibodies. Sequences of the GABARAP-interacting peptides are shown in *black*, and non-interacting peptides are in *gray*. *D*, a schematic map of UBA5. Schematic domain organization and deletion/mutant constructs used in this study are shown. *E*, analysis of UBA5 interaction with UBLs by a GST pulldown assay. Lysates of HEK293 cells expressing GFP or the indicated GFP-UBA5 constructs were precipitated by immobilized GST or GST-UBLs (LC3/GABARAPs and UFM1). Co-precipitated proteins were detected using an anti-GFP antibody. Ponceau S staining of immunoblot membranes shows loading of GST fusion proteins. Representative results from three independent experiments are shown. *WB*, Western blotting. *F*, GST pulldown assay demonstrates interaction of wild-type UBA5 with GABARAPL2, UFM1, and UFC1. LIR/UFIM is required for binding of UBA5 to UFM1 and GABARAPL2, and the C terminus contains a UFC1-binding domain (*UFC1 BD*). *G*, co-immunoprecipitation of GFP-UBA5 and endogenous UFM1 (endogenous GABARAPL2 fails to be detected in the co-immunoprecipitated samples). The indicated forms of human HA-tagged UBA5 were transiently expressed in HEK293 cell, and immunoprecipitated (*IP*) by an anti-HA antibody. Co-precipitated UBLs were detected using the indicated antibodies. Representative results from three independent experiments are shown.

est effect on both UBLs. Additionally, Gly-342 may have a bigger role in the UBA5-UFM1 interaction but less so with the UBA5-GABARAPL2 interaction (Fig. 2*B*).

We then characterized the interaction between UBLs and a synthetic peptide spanning the human UBA5 LIR/UFIM sequence by biophysical methods. The ITC experiment using purified UFM1 protein and the LIR/UFIM peptide showed that the affinity of the UFM1-LIR/UFIM interaction (Fig. 2*C*, *left plot*, and Table 5) lies in the middle range (K_D of 8 μM). On the other hand, the canonical LIR peptide from p62/SQSTM1 (13) showed no binding to UFM1, providing a negative control and

highlighting the specificity of the interaction (Fig. 2*C*, *right plot*). NMR titration experiments confirmed the ITC data for the UFM1-LIR/UFIM interaction. We observed a typical pattern for interaction of two polypeptides with a K_D in the range of 10 μM (Fig. 2*D*). Chemical shift perturbations (CSPs) induced by titration of the nonlabeled LIR/UFIM peptide into ^{15}N-labeled UFM1 were moderate and mostly in intermediate to slow exchange modes. In particular, CSPs of Val-20, Leu-21, Ser-22, Val-23, Ser-26, Ala-31, Val-32, Leu-33, Ala-36, Glu-39, and Ala-63 indicated strong participation of these residues in the intermolecular contacts. However, some resonances of the
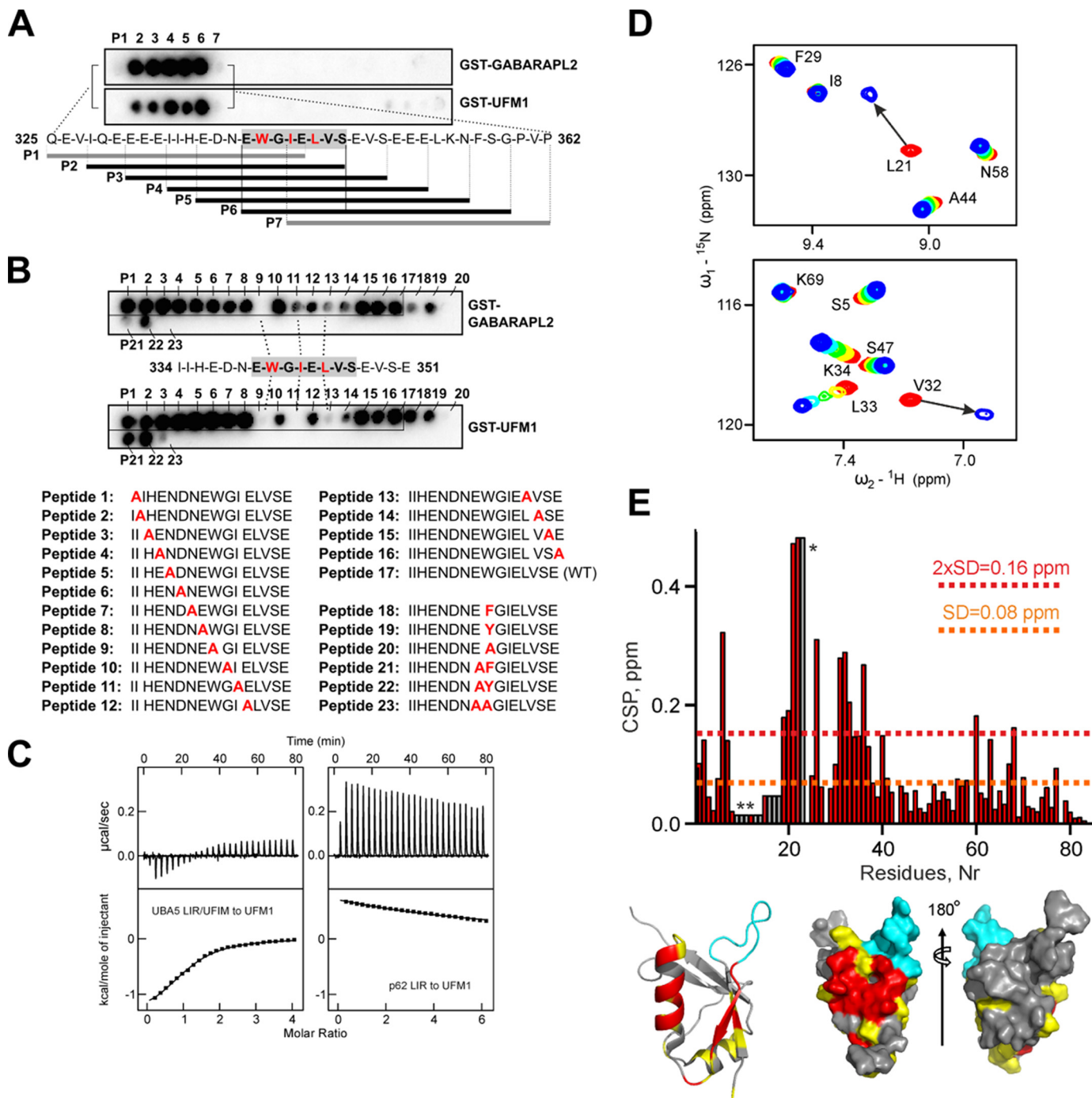
FIGURE 2. LIR/UFIM of UBA5 specifically interacts with UFM1 and LC3/GABARAP proteins. *A*, demonstration of dual specificity of LIR/UFIM for UFM1 and GABARAPL2. Arrays of 20-mer peptides covering the C terminus of UBA5 were synthesized and prepared on PVDF membranes. Each peptide was shifted three amino acids relative to the previous peptide. The arrays were probed with the indicated GST fusion protein, and binding was detected with anti-GST antibodies. Sequences of UBL-interacting peptides are shown in *black*, and non-interacting peptides are in *gray*. *B*, confirmation of the core LIR/UFIM sequence by alanine scanning. PVDF membranes with spotted 20-mer peptides (with indicated alanine substitutions marked in *red* in the peptide sequences) were incubated with the indicated GST-UBLs and probed with an anti-GST antibody to reveal bound proteins. *C*, ITC analysis of the interaction between UBA5 LIR/UFIM and UFM1. A peptide spanning the UBA5 LIR/UFIM sequence (*left panel*) or p62/SQSTM1 LIR as a negative control (*right panel*, without correction of the dilution heat) was titrated to UFM1. The *upper graphs* represent the raw data; the integrated heat of each injection is displayed in the *lower graphs* for each titration. Thermodynamic parameters are summarized in Table 5. *D*, NMR data for interaction between UBA5 LIR/UFIM and UFM1. An overlay of representative areas of the ^{1}H-^{15}N HSQC spectra of ^{15}N-UFM1 to which the nonlabeled UBA5 LIR/UFIM peptide was added stepwise. The *rainbow* color code indicates increasing molar ratios upon titration from free UFM1 (*red*) to saturation (molar ratio 1:2 (*blue*)). *E*, one-dimensional and three-dimensional mapping of CSP induced in UFM1 NMR spectra upon UBA5 LIR/UFIM binding. CSPs calculated for all assigned resonances are shown as *bars*, and the *dashed lines* represent S.D. (1× S.D., *yellow*; 2× S.D., *red*). HN resonances for residues in the loop L1 connecting β1 and β2 (**), as well as V23 in the bound state (*) could not be assigned (*gray bars* showing intensity of the neighbor signals). CSPs were mapped on the UFM1 structure (PDB ID: 1WXS) presented schematically on the *left plot* and as a surface representation in two projections on the *right plot*. Residues that are not affected or are slightly (CSP < 1× S.D.), intermediately (1× S.D. < CSP < 2× S.D.), or strongly (CSP > 2× S.D.) affected by the binding are colored in *gray*, *yellow*, and *red*, respectively. The mostly non-assigned loop L1 is shown in *cyan*.

UFM1 residues show significant CSP in the fast exchange mode (*e.g.* Ser-2, Phe-6, Lys-7, Lys-19, Thr-30, Lys-34, Phe-35, Ala-37, and Phe-40) (Fig. 2*D* and data not shown). They form a cluster at the very N terminus and at the end of α-helix 2 of UFM1, representing the rim of the intermolecular contacts or additional weaker site(s) of interaction. Mapping of CSPs onto the UFM1 sequence and three-dimensional structure shows a specific, well defined interaction site on the UFM1 surface spanning β-strand 2 and α-helix 1 (Fig. 2*E*). The biggest CSPs are localized within the area composed of β-strand 2 and parts of α-helix 1 (residues 20–23 and 31–36), comprising a hydrophobic spot on the UFM1 surface. The UFM1 residues within long loop 1, connecting β-strands 1 and 2, remained mostly unassigned even after formation of the complex with the LIR/UFIM, in agreement with previous observations (34). However, a few assigned residues within loop 1 showed almost no CSPs, indicating a weak contribution of this element to the interaction.

We also analyzed interactions between the LC3/GABARAP proteins and the LIR/UFIM peptide. According to the ITC data, all GABARAP-type proteins had a significant preference for LIR/UFIM binding (Fig. 3*A* and Table 5), with GABARAPL2 showing the highest affinity for LIR/UFIM (K_D of 1.6 μM) followed by GABARAPL1 and GABARAP (with K_D of 6 and 16 μM, respectively). In contrast, LC3 proteins displayed an invariably low affinity (LC3A and LC3C, K_D of ~20 μM each; LC3B, calculated K_D of ~100 μM). NMR titration of the nonlabeled LIR/UFIM peptide to ^{15}N-labeled GABARAPL2 (the strongest interactor) revealed significant CSPs, predominantly in a slow exchange mode, where two positions for each GABARAPL2 resonances, free and bound, are visible in the spectrum at any stage of the titration (Fig. 3*B*, *left panel*). This confirms the relatively high affinity interaction (of a low micromolar range) observed for LIR/UFIM and GABARAPL2 in ITC. Titration of the nonlabeled LIR/UFIM peptide into the ^{15}N-labeled LC3B protein (which binds LIR/UFIM with the lowest affinity) also induced corresponding CSPs, located especially in the LC3B β-strand 2 and adjacent regions (22). However, the exchange regime was fast to intermediate, indicating a much weaker interaction of LC3B with the LIR/UFIM of UBA5 (Fig. 3*B*, *right panel*).

The deviation of the primary sequence of the combined LIR/UFIM in UBA5 from canonical LIRs (Fig. 1*B*) raised a question regarding the structural details of the interaction between this short linear motif with the LC3/GABARAP proteins. To answer this question, we mapped CSPs (induced by titration of the LIR/UFIM peptide in NMR titration experiments) on three-dimensional structures of GABARAPL2 (Fig. 3*C*, *upper panel*) and LC3B (Fig. 3*C*, *lower panel*) and compared those with the well characterized mapping of the LIR found in the selective autophagy receptor optineurin positioned on the surface of LC3B (23) (Fig. 3*C*, *insert*). This comparison revealed that the interaction of the GABARAPL2 and LC3B proteins with UBA5 LIR/UFIM is similar to the canonical LIR-dependent interaction and involves the two hydrophobic pockets HP1 and HP2 and β-strand 2 of the UBL as elements of the intermolecular β-sheet described previously (18, 35). The most perturbed resonances are localized within the corresponding area of HP1 (α-helix 2 and β-strand 2) and HP2 (α-helix 3 and β-strand 2). However, UBA5 LIR/UFIM induced strong CSPs in the α-helix
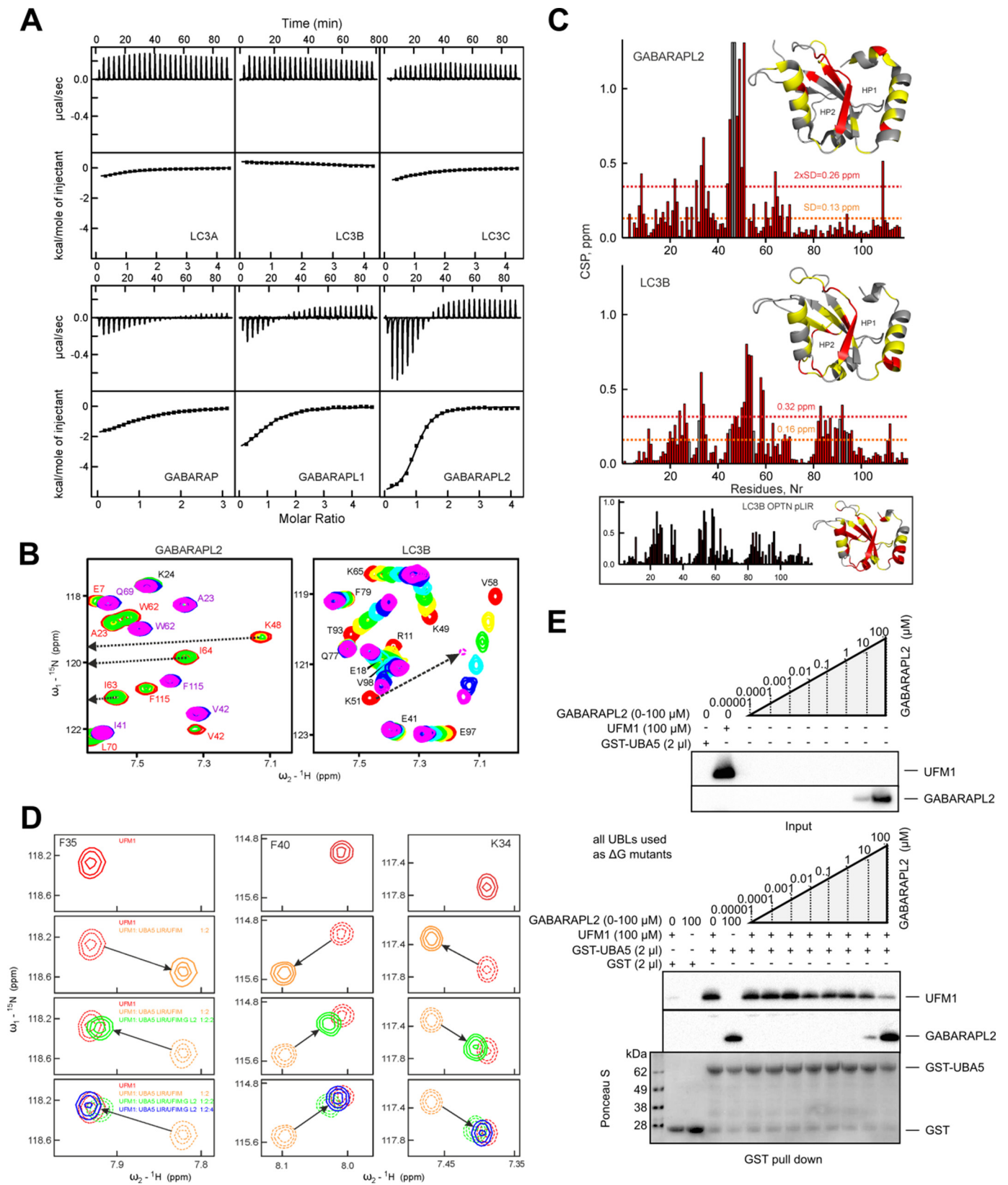
1 (residues Trp-3, Glu-7, Asp-8, and His-9) of GABARAPL2, which are weak for the corresponding residues in LC3B, signifying differences between the LC3 and GABARAP proteins (Fig. 3*C*).

As UFM1 and LC3/GABARAPs bind the same LIR/UFIM peptide, we assessed whether this binding was competitive. We performed an NMR titration analysis of UFM1 in complex with LIR/UFIM in the absence and presence of GABARAPL2, which has the highest affinity to LIR/UFIM. First, we saturated ^{15}N-labeled UFM1 (the backbone HN resonance of Phe-35, directly involved in the intermolecular contacts, is shown in a free form on Fig. 3*D*, *left panel*, *top plot*; two other HN resonances, Phe-40 and Lys-34, are shown on the *right panel* of Fig. 3*D*) with the LIR/UFIM peptide (*second plot* from *top*; molar ratio 1:2) to induce corresponding CSPs in UFM1. Then, we added nonlabeled GABARAPL2 to the same sample at an equimolar ratio to the LIR/UFIM peptide (Fig. 3*D*, *third plot* from *top*; molar ratio 1:2:2). The resonances of UFM1 turned almost quantitatively back to the initial positions, reflecting that almost no free LIR/UFIM peptide remained available in the sample for interactions with UFM1. A further increase in the concentration of nonlabeled GABARAPL2 led to complete stripping of the LIR/UFIM peptide of UFM1 and release of free UFM1 in the sample (Fig. 3*D*, *fourth plot* from *top*; molar ratio 1:2:4). None of the UFM1 HN resonances in the final sample showed positions that would be significantly altered from the positions in the reference spectra, indicating that no ternary complex was formed and that the GABARAPL2·LIR/UFIM complex along with the unbound UFM1 were present. We also performed a competition pull-down experiment in which immobilized GST-UBA5 was pre-bound with purified tag-free UFM1 and subsequently incubated with increasing quantities of purified GABARAPL2. We observed that GABARAPL2 could efficiently outcompete the UFM1-UBA5 interaction, confirming that GABARAPL2 and UFM1 compete for the same binding surface on UBA5 (Fig. 3*E*).

Structure of UBA5 LIR/UFIM in Complex with UFM1— Given the identification of the first UFM1-interacting motif (UFIM), we undertook a detailed structural characterization of the interaction between this short motif and UFM1. We succeeded in co-crystallizing human UFM1 with the LIR/UFIM peptide from human UBA5 and solved the structure of the resulting complex by x-ray crystallography at 2.55 Å resolution (Fig. 4, *A* and *B*, and Table 4). The crystal structure of the UFM1·LIR/UFIM peptide complex was determined by molecular replacement using the NMR structure of the free UFM1 (PDB ID: 1J0G) as the search model. Two UFM1 molecules and one LIR/UFIM peptide (residues [338]DNEWGILV[346] visible in the complex) are present in the asymmetric unit (Fig. 4*A*), with an uneven contribution of the two UFM1 molecules in the interaction with the LIR/UFIM peptide. Although the first UFM1 molecule contributes the most contacts to the peptide and provides the structural basis for specificity in recognition of LIR/UFIM (Fig. 4, *B* and *C*), the second UFM1 molecule may stabilize the resulting complex via contacts with both the LIR/UFIM peptide and the first UFM1 molecule.

The structure of UFM1 in the complex is similar to that determined previously (34) and represents that of Ub-fold pro-

teins. It superimposes well with Ub, SUMO, and LC3 structures with root-mean-square deviations up to 1.8 Å (Fig. 4*E*). Although UFM1 has a fold similar to LC3, it lacks the N-terminal two helices and consequently does not have a hydrophobic pocket corresponding to HP1 in LC3, used for interaction with the aromatic residue in a typical LIR (Fig. 4*E*, *right plot*). A key residue, Ile-343, of the LIR/UFIM occupies the hydrophobic pocket (somewhat similar to HP2 in LC3) formed by the residues of Leu-21, Val-23, Val-32, and Phe-35 of UFM1 (Fig. 4, *B* and *C*). Additionally, Trp-341 and Val-346 of UBA5 LIR/UFIM interact with Pro-28 and Val-20 of UFM1, respectively (Fig. 4*B*). Trp-341 does not intercalate into the hydrophobic core of UFM1 but covers a significant part of the nonpolar surface, playing an important role in the complex stabilization. Leu-345 completes the formation of the hydrophobic cluster within the complex by contacts to the hydrophobic side chains of UFM1 (Val-20, Leu-21, and Val-23) (Fig. 4*B*). Hydrogen bonds between the O and HN atoms of Ser-22 in UFM1 and, correspondingly, the HN and O atoms of Glu-344 in LIR/UFIM, as well as between the O and HN atoms of Val-20 in UFM1 and Val-346 in LIR/UFIM, form an intermolecular antiparallel β-sheet similar to complexes of the SUMO·RANBP2 SIM (SUMO-interacting motif) (PDB ID: 3UIN) and GABARAP·K1 peptides (PDB ID: 3D32) (Fig. 4*D*). A salt bridge between the side chains of Glu-344 in LIR/UFIM and Lys-3 in UFM1 further stabilizes the complex (Fig. 4*B*, *right plot*). In summary, our structural studies reveal that UFM1 interacts with the short linear LIR/UFIM of UBA5 in a way similar to previously described SUMO-SIM and LC3/GABARAP-LIR interactions, with the interesting feature that the LIR/UFIM peptide binds UFM1 in an antiparallel orientation.

Formation of UFM1 Conjugates Depends on Intact LIR/UFIM within UBA5—Finally, we asked what functional significance the LIR/UFIM might have for UBA5 biology. Therefore, we tested whether the formation of UBA5-UFM1 thioester intermediates required for productive ufmylation is affected by mutating the LIR/UFIM sequence in UBA5. Indeed, when expressing wild-type UBA5 in cells, we readily observe a protein band corresponding to the size of the UBA5-UFM1 intermediate. This band runs above UBA5, is immunoreactive with anti-UFM1 antibody, and can be enriched by immunoprecipitation

(Fig. 1*G*). Clearly, disruption of the LIR/UFIM in UBA5 led to strongly reduced formation of the UBA5-UFM1 intermediate (Fig. 1*G*, compare the wild type with W431A/I345A, ΔLIR, and 1–330 aa). This result suggests that the noncovalent interaction between the LIR/UFIM of UBA5 and UFM1 may be critical for either efficient activation of UFM1 or its transfer to the catalytic cysteine of UBA5 (Fig. 1*D*).

To exclude any role for additional cellular proteins in this process, we made use of a previously published *in vitro* thioester formation assay (4) in which the production of thioester intermediates (UBA5-UFM1 and UFC1-UFM1, respectively) reflects the activity of UBA5 as a UFM1-activating (E1) enzyme. In the first reaction, involving purified UBA5 and UFM1, we observed that removal of either the LIR/UFIM or the whole C terminus resulted in reduced amounts of the UBA5-UFM1 intermediate. More subtle disruption of LIR/UFIM, by W341A, G342A, or W341A/L345A mutation, had almost no effect on UBA5-UFM1 conjugate formation *in vitro* (Fig. 5*A*, *left half* of the gel). Interestingly, in the second reaction, also in which purified UFC1 was included together with UBA5 and UFM1, the effect of the LIR/UFIM mutation on the formation of UBA5-UFM1 intermediates was much more pronounced, so that the double mutant W341A/L345A also had a measurable inhibitory effect on the abundance of the UBA5-UFM1 thioester (Fig. 5*A*, *right half* of the gel). W341A and G342A single mutants did not produce any measurable effect under these settings either. The formation of UFC1-UFM1 intermediates was also affected by mutation (W341A/L345A) and removal of LIR/UFIM in UBA5 (Fig. 5*A*, *right half* of the gel). The strongest effect on the UFC1-UFM1 intermediate formation by the UBA5(1–330 aa) mutant, in which the whole C terminus was deleted, could in part be explained by the presence of the UFC1-binding domain (UFC1-BD) within the very C terminus of UBA5 (Ref. 33 and this study (Fig. 1, *A* and *D*)). Importantly, the observed inhibitory effect of LIR/UFIM disruption on the thioester intermediates *in vitro* was rather transient (reaction time of 5 min). Longer reaction times (*i.e.* 30 min) allowed for thioester formation resulted in an almost intact formation of UBA5-UFM1 and UFC1-UFM1 thioester intermediates even in the complete absence of LIR/UFIM (data not shown).

FIGURE 3. **LIR/UFIM of UBA5 specifically interacts with UFM1 and LC3/GABARAP proteins (continued from Fig. 2 legend).** *A*, ITC analysis of the interactions between LIR/UFIM and LC3/GABARAPs. A peptide spanning the LIR/UFIM sequence was titrated to indicated LC3/GABARAP proteins. The *upper graphs* represent the raw data; the integrated heat of each injection is displayed in the *lower graphs* for each titration. Thermodynamic parameters are summarized in Table 5. *B*, NMR data for interaction between UBA5 LIR/UFIM and GABARAPL2 *versus* UBA5 LIR/UFIM and LC3B. An overlay of representative areas of the ^{15}N-^{1}H HSQC spectra of ^{15}N-GABARAPL2 (*left plot*) and ^{15}N-LC3B (*right plot*) to which the nonlabeled UBA5 LIR/UFIM was added stepwise. The *rainbow* color code indicates increasing peptide molar ratios upon titration from free GABARAPL2 and LC3B (*red*) to saturation (*magenta*, 1:2 in the case of GABARAPL2 and 1:4 for LC3B). *C*, three-dimensional modeling of the interaction between UBA5 LIR/UFIM and GABARAPL2 *versus* UBA5 LIR/UFIM and LC3B. CSPs upon interaction with the UBA5 LIR/UFIM were calculated for the GABARAPL2 (*upper plot*) and LC3B (*lower plot*) residues and compared with the well characterized LC3B-optineurin (*OPTN*)-LIR interaction (shown in the *boxed insert*). The *dashed lines* represent S.D. over all resonances, except the three most perturbed ones. CSP mapping of the respective structures (GABARAPL2 (PDB ID: 1EO6) and LC3B (PDB ID: 3VTU)) is shown. Residues that are not or only slightly affected (CSP < 1× S.D.) or intermediately (1× S.D. < CSP < 2× S.D.) or strongly (CSP > 2× S.D.) affected by the UBA5 LIR/UFIM peptide binding are colored in *gray*, *yellow*, and *red*, respectively. *D*, demonstration of UFM1 and GABARAPL2 competition for LIR/UFIM by an NMR competition assay. ^{15}N-^{1}H HSQC spectra of labeled ^{15}N-UFM1 were recorded for the unbound form (*red*) to which the nonlabeled LIR/UFIM peptide from human UBA5 was added to saturation (*yellow*, 1:2 molar ratio). To this complex, nonlabeled GABARAPL2 was added to a molar ratio of 1:2:2 (*green*) and 1:2:4 (*blue*). Competition binding is indicated by the return of all UFM1 resonances to the unbound form due to binding of the LIR/UIFM to GABARAPL2. Well resolved HN resonance of binding-relevant Phe-35 is shown here as a representative example. The *solid lines* represent the visible Phe-35 HN resonance at each stage of the experiment, and the *broken lines* show the positions of Phe-35 HN resonance at previous stages. Additionally, the well resolved HN resonance of binding-relevant residues Phe-40 and Lys-34 are shown. The *solid lines* represent visible Phe-40 and Lys-34 HN resonances at each stage of the experiment, and the *broken lines* show the positions of Phe-40 and Lys-34 HN resonances at previous stages. *E*, demonstration of UFM1 and GABARAPL2 competition for LIR/UFIM by a GST pulldown assay. Immobilized GST-UBA5 with pre-bound UFM1 was incubated with increasing concentrations of GABARAPL2 (0.01 nM–100 μM). Samples were analyzed by immunoblotting using antibodies to the indicated proteins. Ponceau S staining shows the relative loading of GST fusion proteins. Representative results from three independent experiments are shown.
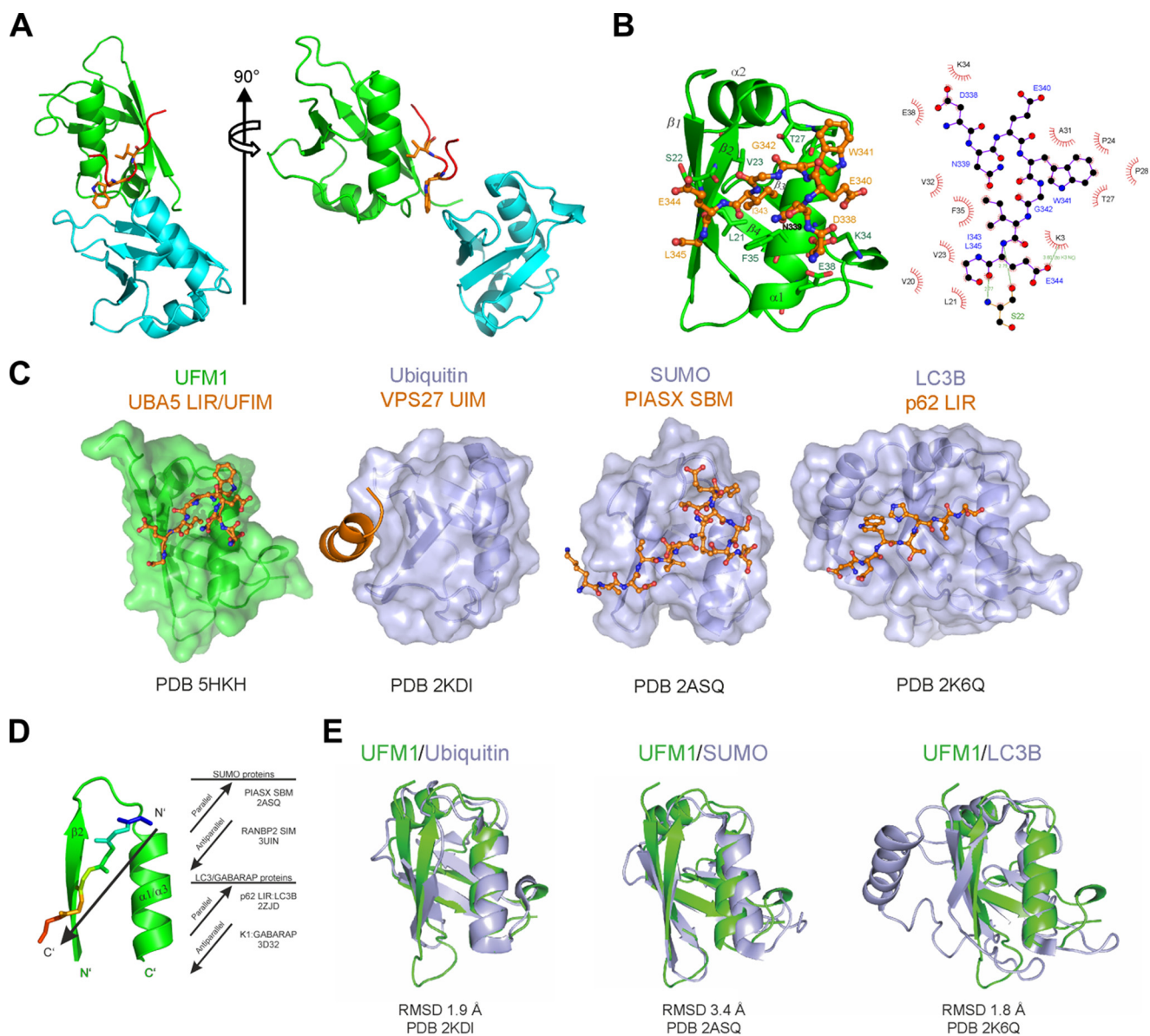
FIGURE 4. **Structure of UBA5 LIR/UFIM in complex with UFM1.** *A*, general structural characteristics of the UBA5 LIR/UFIM·UFM1 complex (PDB ID: 5HKH). The asymmetric unit comprises two UFM1 molecules (*green* and *cyan*; schematic representation) and one UBA5 LIR/UFIM peptide (*orange*; backbone is shown schematically and side chains of Trp-341 and Ile-343 as *sticks*). Although the first UFM1 molecule (*green*) represents the most relevant interactions to LIR/UFIM, the second one stabilizes the complex and could be considered a crystallographic artifact (however, we cannot completely exclude such a dual UFM1 coordination near UBA5 LIR/UFIM). The LIR/UFIM of UBA5 binds to a hydrophobic pocket on the first UFM1 molecule (similar to HP2 for LC3/GABARAP proteins) mainly over the Ile-343 side chain. *B*, specific structural characteristics of the UBA5 LIR/UFIM·UFM1 complex (PDB ID: 5HKH). The *left plot* shows a schematic representation of the UFM1 first molecule (*green*) with the key residues shown as *sticks* (labeled in *green*). The UBA5 LIR/UFIM (*ball-stick* presentation; carbon, oxygen, and nitrogen atoms shown in *orange*, *red*, and *blue*, respectively) binds to a hydrophobic pocket similar to HP2 for LC3/GABARAP proteins. The *right plot* shows a two-dimensional LigPlot diagram with the intermolecular interface between the first UFM1 molecule and the UBA5 LIR/UFIM in a complex. Hydrophobic interactions are represented by *red semicircles* and hydrogen bonds by *green dashed lines*. *C*, comparison of the UFM1·UBA5-LIR/UFIM complex with those of Ub·UIM, SUMO·SBM, and LC3·LIR. Surface and schematic diagrams of UBLs are in the same orientation, with β2 and α1/α3 at the *front*. Interacting peptides (*orange*) are shown for the UBA5 LIR/UFIM·UFM1 complex (PDB ID: 5HKH; first UFM1 molecule is colored *green*), the human Ub in complex with the VPS27 UIM (PDB ID: 2KDI), the human SUMO-1 in complex with the SUMO-binding motif (*SBM*) from PIASX (PDB ID: 2ASQ), and human LC3B in complex with the p62 LIR peptide (PDB ID: 2K6Q; all *light blue*). *D*, schematic representation of the UBA5 LIR/UFIM orientation in complex with UFM1. Typical for the UBL, β-strand β2 and α-helix α1 (α3 in LC3/GABARAP proteins) are shown in the presence of UBA5 LIR/UFIM peptide (backbones are shown as *sticks*; reverse *rainbow* color code: *blue* for the N terminus and *red* for the C terminus). The *arrows* indicate the direction of the peptide (from N to C terminus). Examples of parallel and antiparallel orientations to β2 peptide in known SUMO and LC3/GABARAP proteins are listed. *E*, superimposition of the solved crystal structure of UFM1 (*green*) over those of Ub (*left plot*, *blue*), SUMO1 (*middle plot*), and LC3B (*right plot*). The structures are represented schematically, root-mean-square deviation values for positions of the Cα atoms in each pair are given in Å, and the PDB ID code for each structure is indicated.

Finally, we tested whether conjugation of UFM1 to target proteins in the cell depended on the intact LIR/UFIM in UBA5. We made use of the cellular ufmylation assay in which FLAG-tagged wild-type and LIR/UFIM mutant forms of UBA5 are co-expressed with the C-terminal glycine-exposed form of FLAG-UFM1 as well as Myc-UFL1 and Myc-UFBP1 (the latter serves as a substrate for ufmylation) in HEK293 cells in which *UBA5* was deleted by CRISPR/Cas9 technology (4). UFBP1 is an E3 adaptor protein, dependent on its Lys-267 ufmylation (3); therefore, we used wild-type UFBP1 and its K267R mutant

TABLE 4
Crystallization data collection and refinement statistics
Values in parentheses are for the highest resolution shell. r.m.s.d., root-mean-square deviation.

	Ufm1_Uba5
Data collection statistics	
Beamline	Swiss Light Source PX3
Wavelength (Å)	1.072
Space group	$P\,3_2\,2\,1$
Unit cell (Å)	$a = b = 82.66, c = 62.03$
Resolution (Å)	46.88−2.55 (2.69−2.55)
Observed reflections	160,427 (22,356)
Unique reflections	8,280 (1,177)
Redundancy	19.4 (19.0)
Completeness (%)	100.0 (100.0)
R_{merge}	0.059 (0.800)
$\langle I/\sigma I \rangle$	41.5 (4.4)
Refinement statistics	
Reflections in test set	810
R_{cryst}	19.3
R_{free}	24.8
No. of groups	
Protein residues	173
Ions and ligand atoms	0
Water	25
Wilson B-factor	60.19
r.m.s.d. from ideal geometry	
Bond length (Å)	0.010
Bond angles (°)	1.258
Ramachandran plot statistics	
Favored regions (%)	166 (98.81)
Allowed regions (%)	2 (1.19)
Outliers (%)	0 (0.00)

form as a negative control in this experiment. Upon overexpression of wild-type FLAG-UBA5, or the W341A and G342A single mutants, we observed increased formation of ufmylated UFBP1 species (UFBP1-UFM1, Fig. 5*B*) and the uncharacterized UFM1-protein conjugates (X-UFM1), for which ufmylation in turn depends on ufmylated UFBP1 (Fig. 5*B*). Indeed, UFBP1-UFM1 conjugate formation was abolished in samples expressing UFBP1(K267R) mutants (Fig. 5*B*). In contrast, similar expression levels of LIR/UFIM mutants (W431A/L345A and ΔLIR/UFIM) as well as C-terminal truncation (1−330 aa) of UBA5 resulted in strongly reduced amounts of UFBP1-UFM1 and X-UFM1 conjugates. Equally, the levels of UFC1-UFM1 intermediates were strongly reduced in the samples with W431A/L345A, ΔLIR/UFIM, and 1−330-aa UBA5 mutant expression (UFC1-UFM1, Fig. 5*B*). Intriguingly, despite the aforementioned clear effect on UFM1 conjugates and intermediates, there was no visible effect of LIR/UFIM mutation/deletion on the formation of UBA5-UFM1 intermediates in this assay, likely due to the strong overexpression of both UBA5 and UFM1 (UBA5-UFM1, Fig. 5*B*). We thus concluded that the intact LIR/UFIM is the prerequisite for the full biological activity of UBA5 toward activation and transfer of UFM1 to downstream substrates both *in vitro* and in cells.

Discussion

In this study, we have provided evidence for the presence of a short linear motif in the C terminus of UBA5 that we term LIR/UFIM. It is responsible for the dual binding of UBA5 to distant types of UBLs: UFM1, the cognate UBL for this E1 enzyme; and LC3/GABARAPs, UBLs associated with vesicular transport and autophagy (and so far not shown to be involved in ufmylation). We initially identified LIR/UFIM by its remote

resemblance to the published LIR consensus sequence (W/F/Y)*XX*(L/I/V) (36). However, unlike the canonical LIRs, the combined LIR/UFIM of human UBA5 has up to four core amino acids (EDNE**W**G**IEL**V**S**E), shown in bold, which comprise one aromatic (Trp-341) and three aliphatic (Ile-343, Leu-345, and Val-346) residues and show an unusual spacing (one intermediate residue between the aromatic and the aliphatic amino acids instead of the two in the canonical LIR). Alanine substitutions in the peptide arrays and LIR/UFIM mutations/deletions in the binding assays unequivocally demonstrate the significance of these residues for interactions with both UFM1 and LC3/GABARAPs.

To our knowledge, this is the first report that describes a UFIM in a protein. By solving the crystal structure of the UBA5 UFM1·LIR/UFIM complex, we have provided an insight into the nature of the noncovalent interaction between UBA5 and UFM1 (Fig. 4*B*). Thus, unlike in known LIR-LC3 interactions (*e.g.* (37)), the tryptophan (Trp-341) of LIR/UFIM does not occupy the hydrophobic core of UFM1. Rather, it covers a part of the UFM1 surface and engages in nonpolar interaction with a proline residue, thereby stabilizing the interaction. Ile-343 of LIR/UFIM anchors the interaction by occupying the hydrophobic pocket of UFM1 (which is topologically similar to HP2 of LC3). Leu-345 and Val-346 are both engaged in hydrophobic interactions with the side chains of UFM1. Therefore, the interaction of UBA5 LIR/UFIM with UFM1 shares some features with that of LIR-LC3 but also has unique properties. One of them is also the orientation of the LIR/UFIM in the complex. Short linear motifs reported to date, such as LIR (36), SIM (38, 39), and UFIM (this study), bind UBLs in a roughly similar way by establishing nonpolar interactions between the hydrophobic core of the linear sequence and hydrophobic pockets of the cognate UBL. However, these hydrophobic interactions are context-dependent and can be modulated further via the presence of polar groups, which engage in electrostatic interactions. Consequently, there are two types of LIRs and SIMs, depending on their orientation on LC3 and SUMO. One type binds UBLs in parallel to their β2 (37, 40), whereas the other binds UBLs in an anti-parallel way (40, 41). Here, we have reported the structure of an anti-parallel type of UFM1·LIR/UFIM complex. As this is the first such example, further studies will be required to conclude whether there are also two types of UFIM or if this is the only possible orientation of a UFIM interacting with Ufm1.

The alignment of UBA5 LIR/UFIM with other LIRs may suggest that it is more similar to those found in NDP52 and TAX1BP1, which completely lack an aromatic residue within their LIRs (Fig. 1*B*). Therefore, the role of the aromatic residue (trp-341) in this non-canonical LIR sequence may be to support the dual interaction with UFM1 and LC3/GABARAPs. Interestingly, LIR/UFIM preferentially binds GABARAPs, similar to the LIRs found in NBR1 (14), NIX/BNIP3L (42), KBTBD6/7 (20), ATG13, and ULK1 (19), and ALFY (43). The molecular determinants for this preference are currently not clear. Yet, the increased affinity to the GABARAP subfamily may have a biological significance. Thus, GABARAPs have a comparable affinity to LIR/UFIM as UFM1 (K_D of 8.1 μM for UFM1 *versus* 1.7−14.4 μM for GABARAPs (Table 5)) and may outcompete UFM1 from its complex with UBA5 (Fig. 3, *D* and *E*). These findings suggest a potential role for LC3/GABARAP proteins in
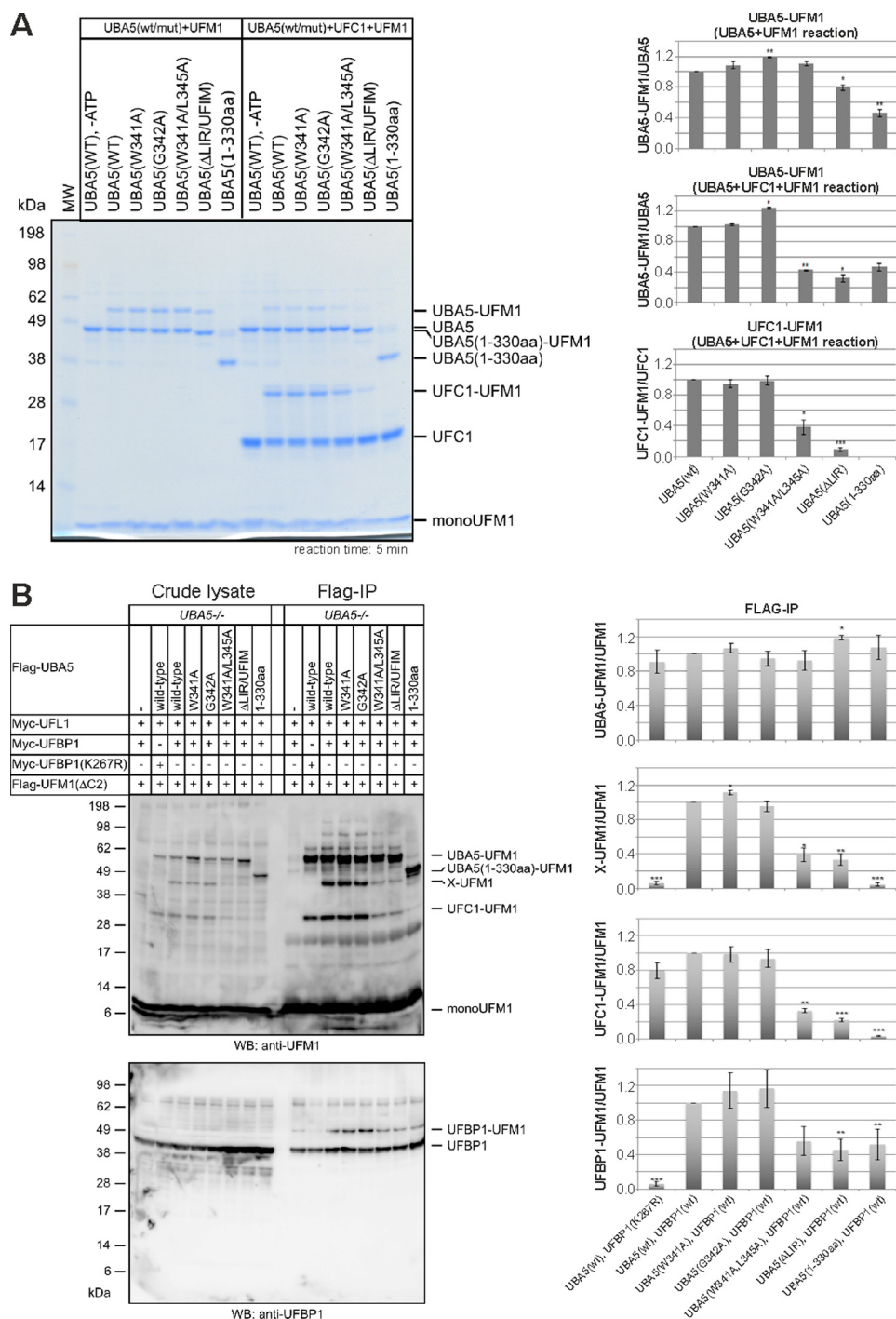
FIGURE 5. **Formation of UFM1 conjugates depends on intact LIR/UFIM within UBA5.** *A*, demonstration of functional significance of LIR/UFIM for UBA5 function by *in vitro* UFM1-UBA5 and UFM1-UFC1 thioester formation assay. Purified UFM1ΔC2 (with exposed C-terminal glycine) and UBA5, or the indicated UBA5 mutants, were incubated with UFC1 (*lanes 1–7*) or without UFC1 (*lanes 8–14*) at 25 °C for 5 min. As a negative control, the reactions were performed in the absence of ATP (*lanes 1* and *8*). The samples were resolved on a nonreducing NuPAGE followed by Coomassie Brilliant Blue staining. Relative ratios of UFM1-UBA5/UBA5 and UFM1-UFC1/UFC1 were determined by the intensity of the bands. Data shown are representative of three independent experiments. *Error bars* represent ± S.E.; *p* values were determined by unpaired *t* test (*, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$). *B*, demonstration of the functional significance of LIR/UFIM for UBA5 function by the cellular ufmylation assay. HEK293 cells, in which *UBA5* was deleted by CRISPR/Cas9 technology, were transfected with the indicated plasmids, and immunoprecipitation experiments (*IP*) with anti-FLAG M2-agarose were performed. FLAG-IP (*right half* of the gel) showed a remarkable reduction of the UFC1-UFM1 intermediate and the UFM1-UFBP1 and X-UFM1 conjugates (where "X" represents an uncharacterized substrate protein) in the samples transfected with UBA5 LIR/UFIM mutants or C-terminal deletion construct (1–330 aa). Relative ratios of UBA5-UFM1/UFM1, X-UFM1/UFM1, UFC1-UFM1/UFM1, and UFBP1-UFM1/UFBP1 (derived by FLAG-IP) were determined by the intensity of the bands. Data shown are representative of three independent experiments. *Error bars* represent ± S.E.; *p* values were determined by unpaired *t* test (*, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$). *WB*, Western blotting.

the regulation of the ufmylation process. Despite this clear possibility, we have so far failed to demonstrate the role of LC3/GABARAP binding for UBA5-mediated UFM1 activation or substrate ufmylation. The presence of GABARAPL2 did not affect the kinetics or abundance of UBA5-UFM1 or UFC1-UFM1 intermediates in the thioester formation assay (data not

TABLE 5
Thermodynamic parameters of the interactions between the UBA5 LIR/UFIM peptide and human UBLs

|	ΔH	ΔS	$-T^*\Delta S$	ΔG	$K_a \times 10^6$	K_d	n
	kcal mol^{-1}	cal mol^{-1} K^{-1}	kcal mol^{-1}	kcal mol^{-1}	M^{-1}	μM	
LC3A	-1.08 ± 0.04	$+18.0$	-5.36	-6.44	0.054 ± 0.005	18.5	1[a]
LC3B	$+3.84 \pm 0.51$	$+30.0$	-8.94	-5.10	0.006 ± 0.001	179	1[a]
LC3C	-1.61 ± 0.02	$+16.0$	-4.77	-6.38	0.048 ± 0.002	20.7	1[a]
GABARAP	-2.51 ± 0.06	$+13.8$	-4.11	-6.62	0.070 ± 0.003	14.4	1.03 ± 0.02
GABARAPL1	-3.53 ± 0.16	$+11.9$	-3.55	-7.08	0.157 ± 0.015	6.4	0.96 ± 0.03
GABARAPL2	-5.91 ± 0.09	$+6.6$	-1.97	-7.88	0.598 ± 0.045	1.7	0.98 ± 0.01
UFM1	-1.51 ± 0.04	$+18.2$	-5.42	-6.93	0.123 ± 0.069	8.1	1.10 ± 0.02

[a] The n value was fixed to 1.00 upon fitting.
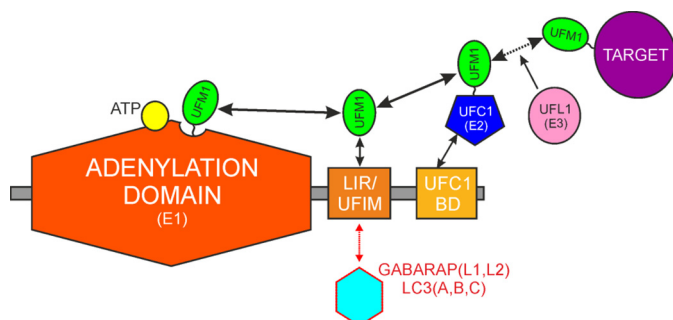
FIGURE 6. **Model for functional interactions between UBA5, UFC1, and UBLs.** The LIR/UFIM motif in the C terminus of UBA5 (E1 enzyme) attracts UFM1. Because of this interaction, UFM1 may occupy the activation site near the ATP-binding site in the catalytic (adenylation) domain with a higher efficiency. LIR/UFIM may further be required for the binding of activated UFM1 for its transfer onto UFC1 (E2 enzyme), which is recruited to UBA5 via its UFC1 binding domain (*UFC1 BD*). UFL1 (E3 enzyme) promotes the transfer of the UFM1 molecule to the substrate protein (target). LC3/GABARAPs might compete UFM1 out of the LIR/UFIM motif. Their function in the ufmylation pathway is presently unknown.

shown). In addition, the physical segregation of UBA5 and LC3/GABARAPs and/or masking of the binding surfaces by post-translational modifications is likely to take place *in vivo*, which further complicates the analysis of the possible cross-talk between the two UBL systems. Our co-immunoprecipitation and co-localization studies failed to detect a strong association between UBA5 and GABARAPL2 (Fig. 1*G* and data not shown). This striking lack of association in cells is in stark contrast to the strong and specific interaction between UBA5 and GABARAPL2 in interaction assays *in vitro* (Fig. 1, *E* and *F*, and Fig. 3*E*). It will be important to study under which conditions UBA5 and LC3/GABARAPs interact in cells and what functional consequences this binding might have for the ufmylation pathway or autophagy.

The LIR/UFIM of UBA5 is required for its role as E1, *i.e.* a UFM1-activating enzyme (Fig. 5, *A* and *B*). Deletion of the C terminus (1–330-aa mutant) in the UBA5 inhibits the formation of UFM1 thioester intermediates and protein conjugates, likely due to the combined lack of UFC1-BD and LIR/UFIM (Fig. 1*D*). UFC1 interaction via the very C-terminal binding motif has been suggested previously by Xie (33) and confirmed by us in this study (Fig. 1*F*). However, deletion of LIR/UFIM alone (which preserves UBA5-UFC1 binding (Fig. 1*F*)) is sufficient to inhibit the formation of UBA5-UFM1 and UFC1-UFM1 (Fig. 5*A*) and UFBP1-UFM1 and X-UFM1 conjugates (Fig. 5*B*). The fact that the longer reaction times for the thioester formation *in vitro* override the dependence of UBA5 on LIR/UFIM (data not shown), and that UFM1 conjugate formation in the cell is not blocked completely when using UBA5

LIR/UFIM mutants (Fig. 5*B*), suggests that the catalytic (adenylation) domain of UBA5 alone (Fig. 1*A*) allows for sufficient, albeit transient, interaction with UFM1, providing compensation for the lack of LIR/UFIM-mediated UFM1 transfer under optimal conditions. Further studies will address the interplay between LIR/UFIM and the catalytic domain of UBA5 and elucidate the relative contribution of the LIR/UFIM to UFM1 activation and transfer along the E1-E2-E3-substrate axis (our current view on that process is summarized in Fig. 6).

In summary, our discovery and characterization of LIR/UFIM and its role in UBA5-UFM1 and UBA5-LC3/GABARAP interactions provide an important insight into the biology of UBA5 and the ufmylation system. Furthermore, the discovery and characterization of the first UFIM will allow identification of further proteins that bind UFM1 noncovalently. By analogy to Ub signaling, in which receptor proteins harboring UBDs (ubiquitin-binding domains) drive cellular events (12, 44), we predict that knowledge of the identity of the UFM1-binding proteins will expedite elucidation of the biological role of this important UBL conjugation pathway.

Author Contributions—S. H., J. H., M. A., D. G. M., and Y. I. designed and performed the experiments. N. R. and F. L. assisted with the ITC and NMR experiments. T. J. designed the peptide array experiment and contributed along with I. D. and V. D. to the concept and manuscript production. M. K. contributed to the concept and directed the research. V. V. R. co-authored the concept, directed research, and designed and performed experiments. V. K. co-authored the concept, directed research, and designed experiments.

References

1. Komatsu, M., Chiba, T., Tatsumi, K., Iemura, S., Tanida, I., Okazaki, N., Ueno, T., Kominami, E., Natsume, T., and Tanaka, K. (2004) A novel protein-conjugating system for Ufm1, a ubiquitin-fold modifier. *EMBO J.* **23,** 1977–1986
2. Kang, S. H., Kim, G. R., Seong, M., Baek, S. H., Seol, J. H., Bang, O. S., Ovaa,

H., Tatsumi, K., Komatsu, M., Tanaka, K., and Chung, C. H. (2007) Two novel ubiquitin-fold modifier 1 (Ufm1)-specific proteases, UfSP1 and UfSP2. *J. Biol. Chem.* **282,** 5256–5262

3. Yoo, H. M., Kang, S. H., Kim, J. Y., Lee, J. E., Seong, M. W., Lee, S. W., Ka, S. H., Sou, Y. S., Komatsu, M., Tanaka, K., Lee, S. T., Noh, D. Y., Baek, S. H., Jeon, Y. J., and Chung, C. H. (2014) Modification of ASC1 by UFM1 is crucial for ERα transactivation and breast cancer development. *Mol. Cell* **56,** 261–274

4. Tatsumi, K., Sou, Y. S., Tada, N., Nakamura, E., Iemura, S., Natsume, T., Kang, S. H., Chung, C. H., Kasahara, M., Kominami, E., Yamamoto, M., Tanaka, K., and Komatsu, M. (2010) A novel type of E3 ligase for the Ufm1 conjugation system. *J. Biol. Chem.* **285,** 5417–5427

5. Zhang, M., Zhu, X., Zhang, Y., Cai, Y., Chen, J., Sivaprakasam, S., Gurav, A., Pi, W., Makala, L., Wu, J., Pace, B., Tuan-Lo, D., Ganapathy, V., Singh, N., and Li, H. (2015) RCAD/Ufl1, a Ufm1 E3 ligase, is essential for hematopoietic stem cell function and murine hematopoiesis. *Cell Death Differ.* **22,** 1922–1934

6. Zhang, Y., Zhang, M., Wu, J., Lei, G., and Li, H. (2012) Transcriptional regulation of the Ufm1 conjugation system in response to disturbance of the endoplasmic reticulum homeostasis and inhibition of vesicle trafficking. *PLoS One* **7,** e48587

7. Lemaire, K., Moura, R. F., Granvik, M., Igoillo-Esteve, M., Hohmeier, H. E., Hendrickx, N., Newgard, C. B., Waelkens, E., Cnop, M., and Schuit, F. (2011) Ubiquitin fold modifier 1 (UFM1) and its target UFBP1 protect pancreatic β cells from ER stress-induced apoptosis. *PLoS One* **6,** e18517

8. Cai, Y., Pi, W., Sivaprakasam, S., Zhu, X., Zhang, M., Chen, J., Makala, L., Lu, C., Wu, J., Teng, Y., Pace, B., Tuan, D., Singh, N., and Li, H. (2015) UFBP1, a key component of the Ufm1 conjugation system, is essential for ufmylation-mediated regulation of erythroid development. *PLoS Genet.* **11,** e1005643

9. Hu, X., Pang, Q., Shen, Q., Liu, H., He, J., Wang, J., Xiong, J., Zhang, H., and Chen, F. (2014) Ubiquitin-fold modifier 1 inhibits apoptosis by suppressing the endoplasmic reticulum stress response in Raw264.7 cells. *Int. J. Mol. Med.* **33,** 1539–1546

10. Tatsumi, K., Yamamoto-Mukai, H., Shimizu, R., Waguri, S., Sou, Y. S., Sakamoto, A., Taya, C., Shitara, H., Hara, T., Chung, C. H., Tanaka, K., Yamamoto, M., and Komatsu, M. (2011) The Ufm1-activating enzyme Uba5 is indispensable for erythroid differentiation in mice. *Nat. Commun.* **2,** 181

11. Hicke, L., Schubert, H. L., and Hill, C. P. (2005) Ubiquitin-binding domains. *Nat. Rev. Mol. Cell Biol.* **6,** 610–621

12. Husnjak, K., and Dikic, I. (2012) Ubiquitin-binding proteins: decoders of ubiquitin-mediated cellular functions. *Annu. Rev. Biochem.* **81,** 291–322

13. Pankiv, S., Clausen, T. H., Lamark, T., Brech, A., Bruun, J. A., Outzen, H., Øvervatn, A., Bjørkøy, G., and Johansen, T. (2007) p62/SQSTM1 binds directly to Atg8/LC3 to facilitate degradation of ubiquitinated protein aggregates by autophagy. *J. Biol. Chem.* **282,** 24131–24145

14. Kirkin, V., Lamark, T., Sou, Y. S., Bjørkøy, G., Nunn, J. L., Bruun, J. A., Shvets, E., McEwan, D. G., Clausen, T. H., Wild, P., Bilusic, I., Theurillat, J. P., Øvervatn, A., Ishii, T., Elazar, Z., Komatsu, M., Dikic, I., and Johansen, T. (2009) A role for NBR1 in autophagosomal degradation of ubiquitinated substrates. *Mol. Cell* **33,** 505–516

15. Nakatogawa, H., Ichimura, Y., and Ohsumi, Y. (2007) Atg8, a ubiquitin-like protein required for autophagosome formation, mediates membrane tethering and hemifusion. *Cell* **130,** 165–178

16. Weidberg, H., Shpilka, T., Shvets, E., Abada, A., Shimron, F., and Elazar, Z. (2011) LC3 and GATE-16 N termini mediate membrane fusion processes required for autophagosome biogenesis. *Dev. Cell* **20,** 444–454

17. Mizushima, N., Yoshimori, T., and Ohsumi, Y. (2011) The role of Atg proteins in autophagosome formation. *Annu. Rev. Cell Dev. Biol.* **27,** 107–132

18. Rogov, V., Dötsch, V., Johansen, T., and Kirkin, V. (2014) Interactions between autophagy receptors and ubiquitin-like proteins form the molecular basis for selective autophagy. *Mol. Cell* **53,** 167–178

19. Alemu, E. A., Lamark, T., Torgersen, K. M., Birgisdottir, A. B., Larsen, K. B., Jain, A., Olsvik, H., Øvervatn, A., Kirkin, V., and Johansen, T. (2012) ATG8 family proteins act as scaffolds for assembly of the ULK complex: sequence requirements for LC3-interacting region (LIR) motifs. *J. Biol.*

Chem. **287,** 39275–39290

20. Genau, H. M., Huber, J., Baschieri, F., Akutsu, M., Dötsch, V., Farhan, H., Rogov, V., and Behrends, C. (2015) CUL3-KBTBD6/KBTBD7 ubiquitin ligase cooperates with GABARAP proteins to spatially restrict TIAM1-RAC1 signaling. *Mol. Cell* **57,** 995–1010

21. Rogov, V. V., Rozenknop, A., Rogova, N. Y., Löhr, F., Tikole, S., Jaravine, V., Güntert, P., Dikic, I., and Dötsch, V. (2012) A universal expression tag for structural and functional studies of proteins. *Chembiochem* **13,** 959–963

22. Rozenknop, A., Rogov, V. V., Rogova, N. Y., Löhr, F., Güntert, P., Dikic, I., and Dötsch, V. (2011) Characterization of the interaction of GABARAPL-1 with the LIR motif of NBR1. *J. Mol. Biol.* **410,** 477–487

23. Rogov, V. V., Suzuki, H., Fiskin, E., Wild, P., Kniss, A., Rozenknop, A., Kato, R., Kawasaki, M., McEwan, D. G., Löhr, F., Güntert, P., Dikic, I., Wakatsuki, S., and Dötsch, V. (2013) Structural basis for phosphorylation-triggered autophagic clearance of *Salmonella. Biochem. J.* **454,** 459–466

24. Salzmann, M., Wider, G., Pervushin, K., Senn, H., and Wüthrich, K. (1999) TROSY-type triple-resonance experiments for sequential NMR assignments of large proteins. *J. Am. Chem. Soc.* **121,** 844–848

25. Grzesiek, S., Anglister, J., and Bax, A. (1993) Correlation of backbone amide and aliphatic side-chain resonances in 13C/15N-enriched proteins by isotropic mixing of 13C magnetization. *J. Magn. Reson. B* **101,** 114–119

26. Logan, T. M., Olejniczak, E. T., Xu, R. X., and Fesik, S. W. (1993) A general method for assigning NMR spectra of denatured proteins using 3D HC-(CO)NH-TOCSY triple resonance experiments. *J. Biomol. NMR* **3,** 225–231

27. Kabsch, W. (2010) XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 125–132

28. Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A., and Wilson, K. S. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67,** 235–242

29. Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 213–221

30. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 486–501

31. Behrends, C., Sowa, M. E., Gygi, S. P., and Harper, J. W. (2010) Network organization of the human autophagy system. *Nature* **466,** 68–76

32. Xie, S. (2014) Characterization, crystallization and preliminary X-ray crystallographic analysis of the Uba5 fragment necessary for high-efficiency activation of Ufm1. *Acta Crystallogr. F Struct. Biol. Commun.* **70,** 765–768

33. Xie, S. (2014) Characterization, crystallization and preliminary X-ray crystallographic analysis of the human Uba5 C-terminus-Ufc1 complex. *Acta Crystallogr. F Struct. Biol. Commun.* **70,** 1093–1097

34. Sasakawa, H., Sakata, E., Yamaguchi, Y., Komatsu, M., Tatsumi, K., Kominami, E., Tanaka, K., and Kato, K. (2006) Solution structure and dynamics of Ufm1, a ubiquitin-fold modifier 1. *Biochem. Biophys. Res. Commun.* **343,** 21–26

35. Noda, N. N., Ohsumi, Y., and Inagaki, F. (2010) Atg8-family interacting motif crucial for selective autophagy. *FEBS Lett.* **584,** 1379–1385

36. Birgisdottir, Å. B., Lamark, T., and Johansen, T. (2013) The LIR motif - crucial for selective autophagy. *J. Cell Sci.* **126,** 3237–3247

37. Ichimura, Y., Kumanomidou, T., Sou, Y. S., Mizushima, T., Ezaki, J., Ueno, T., Kominami, E., Yamane, T., Tanaka, K., and Komatsu, M. (2008) Structural basis for sorting mechanism of p62 in selective autophagy. *J. Biol. Chem.* **283,** 22847–22857

38. Hecker, C. M., Rabiller, M., Haglund, K., Bayer, P., and Dikic, I. (2006) Specification of SUMO1- and SUMO2-interacting motifs. *J. Biol. Chem.* **281,** 16117–16127

39. Kerscher, O. (2007) SUMO junction-what's your function? New insights through SUMO-interacting motifs. *EMBO Rep.* **8,** 550–555

40. Namanja, A. T., Li, Y. J., Su, Y., Wong, S., Lu, J., Colson, L. T., Wu, C., Li, S. S., and Chen, Y. (2012) Insights into high affinity small ubiquitin-like modifier (SUMO) recognition by SUMO-interacting motifs (SIMs) re-

vealed by a combination of NMR and peptide array analysis. *J. Biol. Chem.* **287,** 3231–3240

41. Weiergräber, O. H., Stangler, T., Thielmann, Y., Mohrlüder, J., Wiesehan, K., and Willbold, D. (2008) Ligand binding mode of GABAA receptor-associated protein. *J. Mol. Biol.* **381,** 1320–1331

42. Novak, I., Kirkin, V., McEwan, D. G., Zhang, J., Wild, P., Rozenknop, A., Rogov, V., Löhr, F., Popovic, D., Occhipinti, A., Reichert, A. S., Terzic, J., Dötsch, V., Ney, P. A., and Dikic, I. (2010) Nix is a selective autophagy receptor for mitochondrial clearance. *EMBO Rep.* **11,** 45–51

43. Lystad, A. H., Ichimura, Y., Takagi, K., Yang, Y., Pankiv, S., Kanegae, Y., Kageyama, S., Suzuki, M., Saito, I., Mizushima, T., Komatsu, M., and Simonsen, A. (2014) Structural determinants in GABARAP required for the selective binding and recruitment of ALFY to LC3B-positive structures. *EMBO Rep.* **15,** 557–565

44. Kirkin, V., and Dikic, I. (2007) Role of ubiquitin- and Ubl-binding proteins in cell signaling. *Curr. Opin. Cell Biol.* **19,** 199–205