AF2-multimer guided high accuracy prediction of typical and atypical ATG8 binding motifs

Tarhan Ibrahim[1], Virendrasinh Khandare[1,2], Federico Gabriel Mirkin[1,3], Yasin Tumtas[1], Doryen Bubeck[1*] and Tolga O. Bozkurt[1*]

[1] Department of Life Sciences, Imperial College London, London SW7 2BX, United Kingdom

[2] Department of Agrotechnology and Food Sciences, Biochemistry, Wageningen University and Research, 6708 PB Wageningen, Netherlands

[3] INGEBI-CONICET, Ciudad Autonoma de Buenos Aires, Argentina

* Corresponding authors
E-mail: o.bozkurt@imperial.ac.uk
E-mail: d.bubeck@imperial.ac.uk

Abstract

Macroautophagy/autophagy is an intracellular degradation process central to cellular homeostasis and defense against pathogens in eukaryotic cells. Regulation of autophagy relies on hierarchical binding of autophagy cargo receptors and adaptors to ATG8/LC3 protein family members. Interactions with ATG8/LC3 are typically facilitated by a conserved, short linear sequence, referred to as the ATG8/LC3 interacting motif/region (AIM/LIR), present in autophagy adaptors and receptors as well as pathogen virulence factors targeting host autophagy machinery. Since the canonical AIM/LIR sequence can be found in many proteins, identifying functional AIM/LIR motifs has proven challenging. Here we show that protein modelling using Alphafold-Multimer (AF2-multimer) identifies both canonical and atypical AIM/LIR motifs with a high level of accuracy. AF2-multimer can be modified to detect additional functional AIM/LIR motifs by using protein sequences with mutations in primary AIM/LIR residues. By combining protein modelling data from AF2-multimer with phylogenetic analysis of protein sequences and protein-protein interaction assays, we demonstrate that AF2-multimer predicts the physiologically relevant AIM motif in the ATG8-interacting protein 2 (ATI-2) as well as the previously uncharacterized non-canonical AIM motif in ATG3 from potato (*Solanum tuberosum*). AF2-multimer also identified the AIM/LIR motifs in pathogen-encoded virulence factors that target ATG8 members in their plant and human hosts, revealing that cross-kingdom ATG8-LIR/AIM associations can also be predicted by AF2-multimer. We conclude that the AF2-guided discovery of autophagy adaptors/receptors will substantially accelerate our understanding of the molecular basis of autophagy in all biological kingdoms.

2

Introduction

To withstand stressful conditions, eukaryotic cells employ a fundamental intracellular catabolic process known as macroautophagy or autophagy. Autophagy is central to cellular homeostasis in all eukaryotes, from humans to yeasts and plants since it mediates adaptation to harmful environmental conditions by eliminating damaged and toxic cellular components as well as invading microbes. Autophagy plays critical roles in various physiological and pathological conditions, particularly defence against pathogens, thus directly affecting plant and human health [1].

Autophagy is a multistep process initiated by the induction of an isolation membrane that expands and closes to form a double-membraned vesicle named the autophagosome. Autophagy cargoes are typically loaded into the inner leaflet of the isolation membrane [2]. Mature autophagosomes are transported to fuse with lysosomes, which in turn digest the captured cargoes [3]. Alternatively, some autophagosomes are re-routed towards the cell surface to discharge cytoplasmic components outside the cell, through a process called secretory autophagy [4]. In yeast (*Saccharomyces cerevisiae*), autophagy is coordinated by more than thirty core proteins known as the ATG (autophagy-related) proteins that are recruited to autophagosome formation sites in a hierarchical manner [2]. Many of these proteins are also conserved in plants and humans, such as the ubiquitin-like ATG8-family proteins that play central roles in virtually all steps of autophagy, from cargo sequestration to transport and lysosomal fusion of autophagosomes [5, 6]. Although yeast has only a single copy of ATG8, the protein has diversified into multiple isoforms in plants (ATG8A–I) and humans (referred to as the LC3/GABARAP family), forming family-specific ATG8 clades [7]. ATG8 associates with proteins that regulate autophagy initiation, such as ATG7, ATG3, and ATG1 [8, 9], and decorates both inner and outer autophagosomal membranes to coordinate autophagosome biogenesis and transport. ATG8 sequesters specific cargoes inside the autophagosomes by interacting with autophagy cargo receptors that capture specific cargoes [10, 11]. In addition, ATG8 binds to autophagy adaptors on the autophagosome surface to modulate autophagosome transport and fusion events [12, 13].

ATG8-interacting proteins carry a short linear motif (SLIM) called the ATG8/LC3 interaction motif/region (AIM/LIR) that interacts with the core autophagy machinery as well as autophagy adaptors and receptors [14]. The core AIM/LIR motif sequence ([W,Y,F][X][X][L,I,V]) consists of an aromatic amino acid followed by any two amino acids and a hydrophobic residue. The first aromatic and last hydrophobic amino acids of the core AIM/LIR motif bind to two hydrophobic pockets — also known as W (hydrophobic pocket 1) and L (hydrophobic pocket 2) pockets — on the surface of a ubiquitin-like fold in ATG8 proteins. The sequences flanking the aromatic residue of the AIM/LIR typically consist of negatively charged residues, which enhance AIM/LIR docking by forming polar

3

interactions with the positively charged residues surrounding the W and L pockets [15]. ATG8 binding can be improved further via post-translational modifications, such as phosphorylation of the residues flanking the core AIM/LIR regions [16, 17]. However, non-canonical AIM/LIR motifs that mediate ATG8 binding have also been discovered, expanding the spectrum of residues that facilitate ATG8 association [18]. The emerging paradigm is that autophagy is primarily orchestrated through sequential binding of core autophagy components, autophagy adaptors, and cargo receptors to the ATG8 family members via canonical/non-canonical AIM/LIR motifs. Therefore, identifying and characterizing the AIM/LIR residues has been a crucial step in dissecting the molecular basis of autophagy regulation from initiation to selective cargo sorting and autophagosome transport.

As the canonical AIM/LIR sequence consensus occurs in many proteins, identifying functional AIM/LIR motifs has proven challenging. Thus, a simple search for the [W,Y,F][X][X][L,I,V] amino acid sequence predicts many false AIM/LIR motifs. Since proteins of interest might also have multiple potential AIM/LIR motifs that match the consensus pattern, these motifs must be validated through mutagenesis, peptide binding, and autophagy assays. For instance, the yeast ATG7 has 16 residues that match the core AIM/LIR consensus sequence, yet none have been functionally validated. Likewise, the human NDP52 autophagy cargo receptor has 8 predicted AIM/LIRs, but binds to LC3 through a different, non-canonical LIR [18]. However, a distinctive feature of the AIM/LIR motif that can help narrow down the number of candidates has been discovered: functionally validated AIM/LIR motifs are typically located in intrinsically disordered protein regions (IDPRs) [14]. To help discover functional AIM/LIRs, *in silico* AIM/LIR prediction tools such as iLIR [19], hfAIM [20], and pLIRm [21] can also be used. Although these computational tools can be used somewhat successfully to predict canonical AIM/LIR motifs, none can detect non-canonical AIM/LIR motifs. Additionally, none of these methods determine the spatial distribution of AIM/LIR motifs or flanking residues on ATG8 proteins; however, this level of resolution combined with the ability to determine side chain interactions established by the flanking AIM/LIR residues would significantly improve investigations into ATG8 binding and help determine the ATG8 binding specificity of various autophagy-related proteins. Furthermore, these tools would accelerate studies aiming to dissect the evolutionary dynamics of autophagy regulation in eukaryotes.

The single-chain protein structure prediction tool AlphaFold2 (AF2) [22], and the recently retrained AlphaFold-Multimer (AF2-multimer) system that can predict homomeric and heteromeric interfaces [23], has sparked a renewed interest in applying structural biology to understanding complex cellular processes. In this study, we investigated whether AF2-multimer could identify ATG8 binding structures mediated by AIM/LIR motifs carried by various proteins that have been shown to bind ATG8. AF2-multimer showed 90% accuracy in determining AIM/LIR motifs in 33 experimentally validated proteins that carry functional AIM/LIR motifs. At present, AF2-multimer can identify multiple

4

AIM/LIR residues even after *in silico* deletion/mutagenesis of the primary AIM/LIR motif. Strikingly, AF2-multimer predicted all three non-canonical AIM/LIR motifs that are experimentally validated, as well as the previously uncharacterized AIM motif in plant ATG3, which is not possible with other current prediction tools. Furthermore, AF2-multimer predicted functional AIM/LIRs in 3 of the 4 tested proteins encoded by plant and human pathogens which target host ATG8 proteins, indicating that AF2-multimer can also detect cross-kingdom AIM/LIR–ATG8 interactions. Our study highlights the potential of AF2-multimer in identifying AIM/LIR residues and provides a framework for discovery of new autophagy receptors and adaptors. This AI-guided approach will substantially accelerate our understanding of the molecular basis of autophagy in all kingdoms of life.

Results

AF2-multimer predicts experimentally validated canonical AIM/LIR motifs with high accuracy.

To determine the extent to which AF2-multimer predicts AIM/LIR motifs, we decided to run AF2-multimer using a subset of proteins that carry experimentally validated, canonical (W/Y/ F-X-X-L/I/V) AIM/LIR motifs along with their corresponding ATG8 proteins from plants, humans, and yeast. The AF2-predicted models of plant ATG8CL, human GABARAP, and yeast ATG8 proteins display highly similar backbone and side-chain conformations to previously solved crystal structures [24-26]. All tested ATG8 AF2-models have root mean square deviation (RMSD) scores close to 0.5Å (*Sc*ATG8: 0.51Å; *Hs*GABARAP: 0.51Å; *St*ATG8CL: 0.50Å), suggesting AF2-multimer can be used to identify ATG8 binding proteins. Reflecting the conserved nature of the AIM/LIR consensus sequence, amino acid alignments between yeast, plant, and human ATG8 isoforms showed that amino acids forming the W- (site one) and L- (site two) pockets, which accommodate the AIM/LIR motifs, show high sequence similarity (Figure 1A/B). This observation suggests that universal binding patterns govern ATG8–LIR/AIM associations, and that these patterns can be predicted by AF2-multimer. Among the 36 tested proteins with experimentally validated canonical ([W,Y,F][X][X][L,I,V]) AIM/LIR motifs, AF2-multimer predicted 33 (~90%) of them accurately (Supplemental Table 1, Figure 1C-H, Supplemental Figure S1A-F). Consistent with previous work that revealed canonical AIM/LIRs are typically found in structurally disordered regions [14], nearly all AF2-multimer predicted AIM/LIRs appeared in IDPRs, while a few were predicted to be partially disordered. These findings are consistent with the view that AIM/LIR motifs are structurally flexible protein interaction interfaces, referred to as short linear motifs (SLiMs), that display dynamic structural plasticity to precisely associate with their substrates [14].
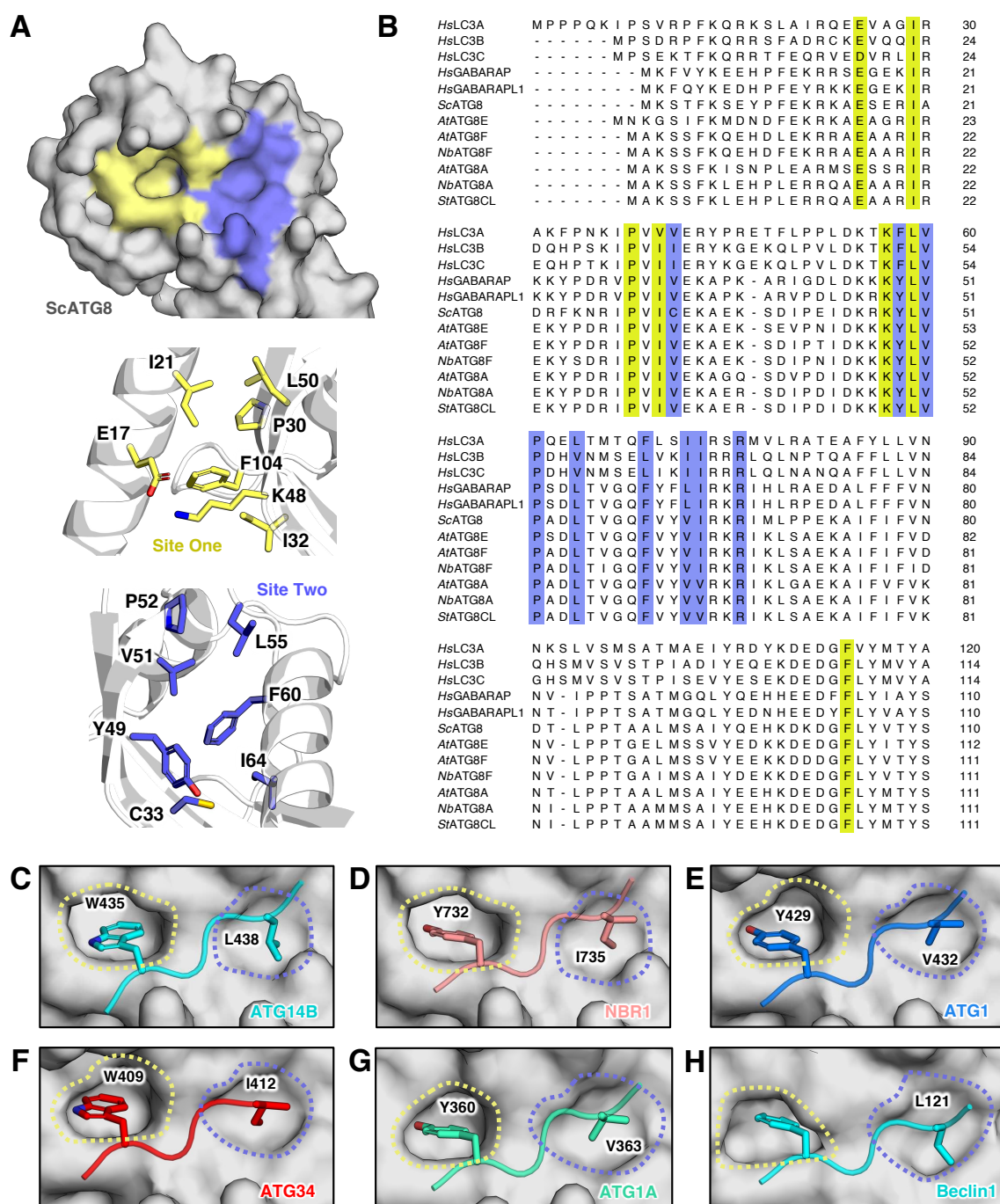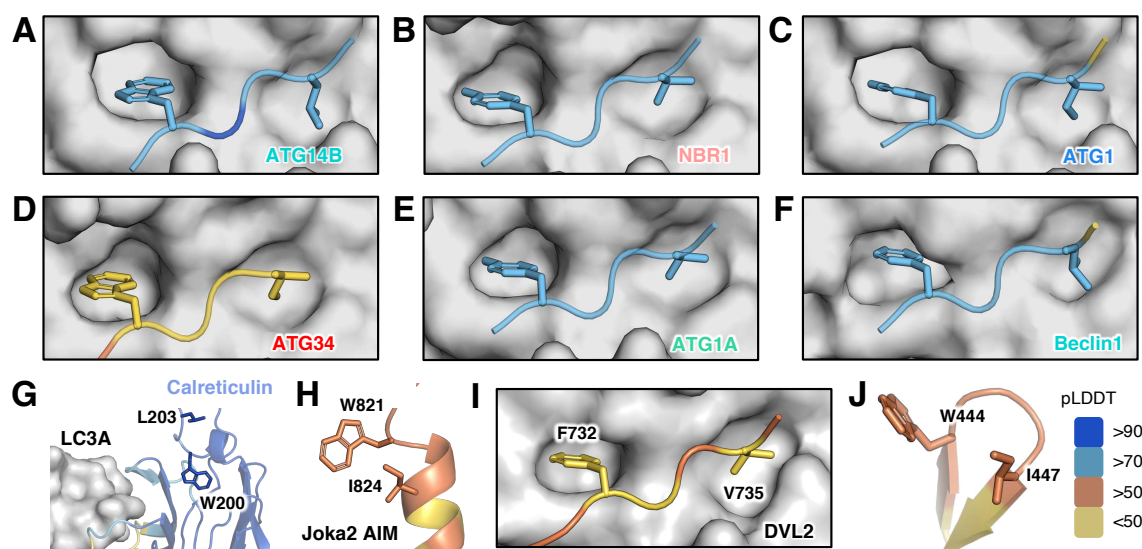
5

Figure 1. AF2-multimer can correctly identify experimentally validated canonical AIM/LIRs in ATG8-interacting proteins. (A) The yeast ATG8 (surface representation – grey) has two AIM/LIR pockets, known as W-site (site one – yellow) and L-site (site two – slate blue) (top panel). The core of the pockets is formed by residues phenylalanine (F), leucine (L), valine (V), isoleucine (I), tyrosine (Y), proline (P) and polar amino acids glutamate (E), lysine (K) and cysteine (C) guard the pocket. Residues comprising site one and site two are shown in the middle and bottom panels, respectively. **(B)** These residues are conserved amongst the ATG8-family of all species used for this study. Residues conserved and forming the W- (yellow) and L- (slate blue) sites have been highlighted accordingly. Some examples of accurate AF2 multimer validation of experimental work include human GABARAP interactions with **(C)** ATG14B (cyan) and **(D)** NBR1 (salmon), yeast ATG8 interactions with **(E)** ATG1 (blue) and **(F)** ATG34 (red), **(G)** *A. thaliana* ATG8A with ATG1A (green) and **(H)** *N. benthamiana* ATG8A with Beclin1 (cyan). Residues interacting with the AIM/LIR pockets identified by AF2 multimer and experimentation have been highlighted via their one letter amino acid code and position within the protein and shown with their side chains. Side chains have been coloured based on elements by PyMOL, where oxygen is red, and nitrogen is blue.

Supplemental Figure 1. AF2-multimer calculated confidence for prediction of experimentally validated canonical AIM/LIRs. (A-F) AF2 and experimentally verified canonical AIM/LIR ATG8-family interactions, shown in Figure 1 C-H, coloured based on AF2-calculated prediction score; predicted local distance difference test (pLDDT). Blue indicates regions of a protein with a score of over 90, meaning a very high confidence prediction. Scores between 70-90 are represented with light blue, which accounts for a high confidence score. Scores between 50-70 are considered low (orange) and anything below 50 is a very low confidence prediction (yellow). **(G)** Calreticulin does not interact with LC3A via its canonical LIR, which seems to be positioned away from LC3A (surface representation – grey). However, Calreticulin may interact with the C-terminal region of LC3A. **(H)** Joka2 AIM residue ILE824 is localised in the alpha-helix and no ATG8 interaction is observed. **(I)** DVL2 may interact with LC3A LIR pockets via a different canonical AIM, **(J)** as the experimentally studied LIR was observed away from LC3A and structured into a beta-sheet. Calreticulin, Joka2 and DVL2 are coloured based on the AF2-calculated prediction confidence score, pLDDT.

Supplemental Table 1. ATG8-interacting proteins used in this study.

Despite its high accuracy, AF2-multimer failed to detect AIM/LIR motifs in plant Joka2 (also known as NBR1) and human Calreticulin proteins [27, 28]. The confidence scores determined by the predicted local distance difference test (pLDDT) [22] in Joka2's AIM motif were low (~50) (Supplemental Figure S1H), which could account for the failure of AIM/LIR prediction. Although confidence scores in the previously identified LIR motif of Calreticulin were high (>70), AF2-multimer did not identify any interaction between LIR and LIR docking (LDS) regions (Supplemental Figure S1G). Superimposition of the Calreticulin crystal structure (3POW) [29] with the AF2-multimer model showed a high correlation (RMSD: 0.33Å), indicating that AF2-multimer's inability to identify the functional LIR is not a consequence of structural mis-prediction. Of note, the AIM/LIR motifs in Joka2 and Calreticulin are embedded in short, disordered regions (5 and 6 amino acids long, respectively) that are flanked by alpha-helices and beta-strands, respectively. The flexibility of the AIM/LIR motifs may be constrained by these ordered regions, which could hinder ATG8 docking. We speculate that AF2-multimer predictions might fail in such cases because the AIM/LIR–ATG8 association is

7

somewhat conditional, since post-translational modifications, intramolecular interactions, or interactions with other proteins can influence ATG8 binding [16, 30, 31].

Intriguingly, AF2-multimer revealed a distinct canonical LIR motif (FVDV) in the human autophagy substrate DVL2 that differed from the one previously discovered (WLKI) (Supplemental Figure S1I). Mutating DVL2^{WLKI} to DVL2^{ALKA} reduced, but did not prevent, LC3 binding [32], indicating that additional LIR residues, as predicted by AF2-multimer, may influence LC3 binding. The new LIR predicted by AF2-multimer (FVDV) is in the C-terminal IDPR, as opposed to the previously verified LIR motif (WLKI), which is situated between two beta-strands (Supplemental Figure S1J). This may account for AF2-multimer's preference for the second LIR. Nevertheless, further research is needed to determine whether the additional LIR predicted by AF2-multimer is valid. Thus, AF2-multimer is a powerful tool for predicting canonical AIM/LIR motifs with a high level of accuracy; however, several limitations, such as conditional ATG8 binding due to post-translational modifications, may limit its effectiveness.

AF2-multimer can predict multiple AIM/LIR motifs.

The human E3 ubiquitin ligase NEDD4 was previously reported to carry two LIR motifs that bind LC3B. Deletion analysis of NEDD4 revealed that LIR1 (WVVL) is not necessary for LC3 binding, whereas deletion of both LIR1 and LIR2 (WEII) impairs LC3 binding [33]. Although LIR2 is more critical for LC3 binding, it is possible that LIR1 is also functional. Interestingly, AF2-multimer predicted LIR1 over LIR2 for AIM pocket association (Figure 2A). To determine whether AF2-multimer could also predict LIR2, we truncated LIR1 and re-ran AF2-multimer prediction [33]. In line with previous findings, LIR2 displayed a perfect association with the AIM pocket in the absence of LIR1 (Figure 2B). By contrast, an *in silico* truncated NEDD4 sequence lacking both LIRs failed to form any association with the AIM pocket on LC3B (Supplemental Figure S2A). Consistent with our AF2-predictions, a different study showed that both LIR1 and LIR2 peptides bind LC3B with similar KDs [34]. Phylogenetic analysis revealed that both LIRs are conserved, further supporting the AF2-multimer predictions that NEDD4 has two functional LIR motifs (Figure 2C). To determine whether AF2-multimer predicts LIR1 over LIR2 simply because LIR1 is positioned upstream of LIR2, we swapped the position of LIR2 (X_4WEEIX_4) with that of LIR1 (X_4WVVLX_4) and re-ran AF2-multimer. After this change was made, AF2-multimer predicted that LIR2 was the main functional LIR motif (Supplemental Figure S2B), indicating that AF2 prioritises the first functional AIM/LIR.
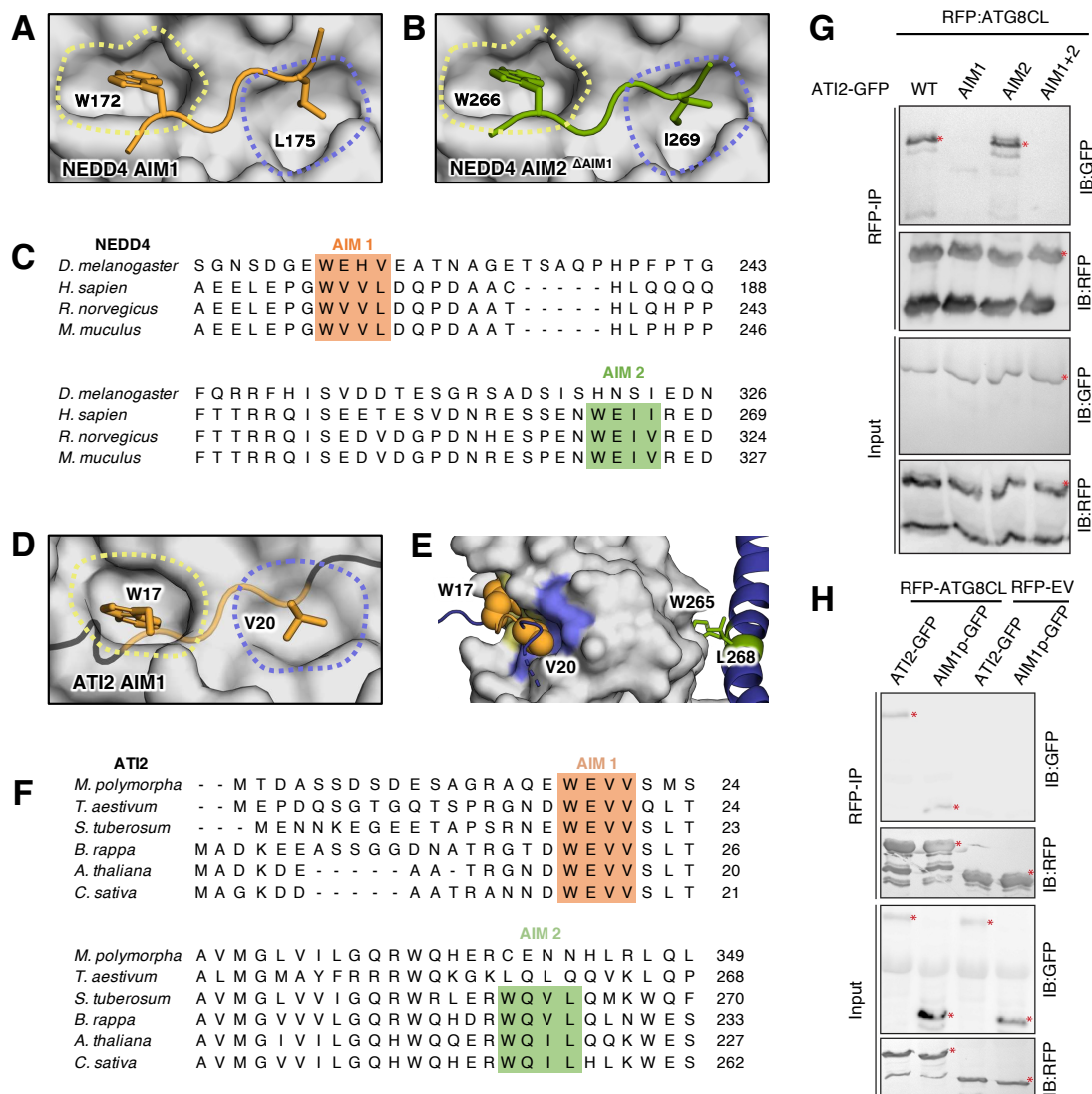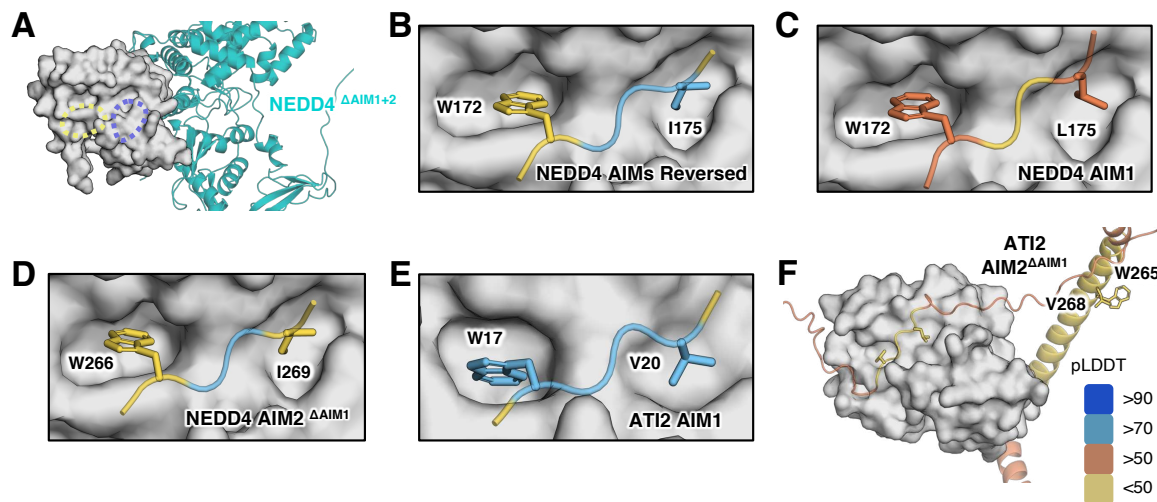
Figure 2. AF2 multimer can correctly identify a functional AIM/LIR amongst various targets. (A) AF2-multimer predicts AIM1 (172)WxxL(175) of NEDD4 as the interactor for LC3B pockets. **(B)** Upon the truncation of AIM1, AF2-multimer predicts another interface at (266)WxxI(269), which was identified as AIM2, previously. **(C)** Two AIMs of NEDD4 are conserved among *Drosophila melanogaster*, *Homo sapiens*, *Rattus norvegicus* and *Mus musculus*. **(D)** ATI2 contains two-functional AIMs and AF2-multimer predicts the interface at AIM1 (17)WxxV(20) as the functional AIM. **(E)** AIM2 of ATI2, (265)WxxV(268), is localised in an alpha-helix and a distance away from the LC3B AIM pockets. **(F)** AIM1 of ATI2 is conserved among *Marchantia polymorpha*, *Triticum aestivum*, *Solanum tuberosum*, *Brassica rapa*, *Arabidopsis thaliana* and *Cannabis sativa*, whereas AIM2 is not conserved in *Marchantia polymorpha* and *Triticum aestivum*. AIM1 and AIM2 for both NEDD4 and ATI2 have been coloured as orange and green, respectively. (G) In planta co-immunoprecipitation between ATG8CL and ATI2 and its mutants (ATI2AIM1, ATI2AIM2 or ATI2AIM1+2). RFP-ATG8CL was transiently co-expressed with either ATI2-GFP, ATI2AIM1GFP, ATI2AIM2-GFP or ATI2AIM1+2-GFP. Red asterisks indicate expected band sizes. (H) In planta co-immunoprecipitation between ATG8CL and ATI2 and ATI2 AIM1 peptide. RFP-ATG8CL or RFP-EV control was transiently co-expressed with either ATI2-GFP or ATI2 AIM1p-GFP. Red asterisks indicate expected band sizes.

Supplemental Figure 2. AF2-multimer can distinguish between functional AIM/LIRs of NEDD4 and ATI2. **(A)** The pockets of LC3B remain empty when both LIRs of NEDD4 are truncated. **(B)** Upon the reversal of LIR -1 and -2 positions, AF2-multimer favours LIR1 (previously LIR2). **(C-D)** AF2-multimer prediction confidence of the NEDD4 AIM1 and AIM2 in complex with LC3B. **(E)** AF2-multimer prediction confidence for ATI2 AIM1 in complex with ATG8CL. **(F)** ATI2 AIM2 cannot occupy ATG8CL pockets when AIM1 is truncated. NEDD4 and ATI2 in panels B-F are coloured based on the AF2-calculated prediction confidence score; predicted local distance difference test (pLDDT). Blue indicates regions of a protein with a score of over 90, meaning a very high confidence prediction. Scores between 70-90 are represented with light blue, which accounts for a high confidence score. Scores between 50-70 are considered low (orange) and anything below 50 is a very low confidence prediction (yellow).

We next used this approach to assess the plant ATG8 interacting protein-2 (ATI-2), which has two predicted AIM motifs [35]. A recent yeast-two-hybrid analysis of ATI-2 and ATG8f proteins from the model plant Arabidopsis (*Arabidopsis thaliana*) indicated that the N-terminal AIM1 in ATI-2 is necessary and sufficient for ATG8 interaction [36]. We examined whether AF2-multimer could accurately predict the functional AIM in the potato (*Solanum tuberosum*) ATI-2 protein. Consistent with the experimental evidence provided for the Arabidopsis ATI-2 homologue [36], AF2-multimer identified AIM1 of the potato ATI-2 as associating with the AIM pocket on the ATG8CL variant (Figure 2D/E). *In silico* truncation of the core AIM1 residues ablated any AIM2–ATG8 association as predicted by AF2-multimer (Supplemental Figure S2F). Multiple sequence alignments revealed that although the potato AIM1 sequence is conserved, the AIM2 sequence is only conserved in closely related dicot plants but not in distantly related plant species, suggesting that AIM2 may not be a functional AIM (Figure 2F). To validate this AIM prediction, we tested the interaction of potato ATG8CL with ATI-2 constructs harbouring AIM1, AIM2, and AIM1,2 double mutations in plant cells. In agreement with the AF2-multimer predictions, our co-immunoprecipitation experiments showed that ATI-2[AIM1] and ATI-2[AIM1,2] double mutants fail to interact with ATG8CL, whereas the ATI-2[AIM2] mutant binds ATG8CL much like the wild-type ATI-2 protein (Figure 2G). We further validated these findings by generating an AIM1 peptide construct that effectively bound ATG8CL *in planta* (Figure 2H). These results indicate that while AF2-multimer prioritises the first functional AIM/LIR in a protein

10

containing several functional AIM/LIR motifs, it prevents the prediction of spurious AIM/LIRs. Thus, an *in-silico* protein truncation approach can be used to identify multiple AIM/LIRs using AF2-multimer.

AF2-multimer can predict non-canonical AIM/LIRs

Since there are no established techniques to predict non-canonical AIM/LIR motifs, their discovery poses a challenge [37]. We investigated the ability of AF2-multimer to recognise previously characterised non-canonical LIR motifs in the human autophagy cargo receptors NDP52 and TAX1BP1 [18, 38]. AF2-multimer accurately identified the validated non-canonical LIR motifs of NDP52 (ILVV) and TAX1BP1 (MLVV) with high confidence. In these two resolved models, I/M residues occupy the W pocket (site one), while the final Vs are located in the L pocket (site two) of their respective LC3 proteins. Central L and V residues, however, appear to form additional hydrophobic surface interactions (Figure 3A/B). We also used AF2-multimer to predict the yeast ER-phagy cargo receptor ATG40, which carries a non-canonical AIM (YDFM) with an oddly positioned M in the last AIM position rather than the standard L/I/V residues [39]. AF2-multimer accurately predicted the non-canonical AIM in this case as well and showed that the atypical M residue occupies the L pocket on ATG8 as expected.
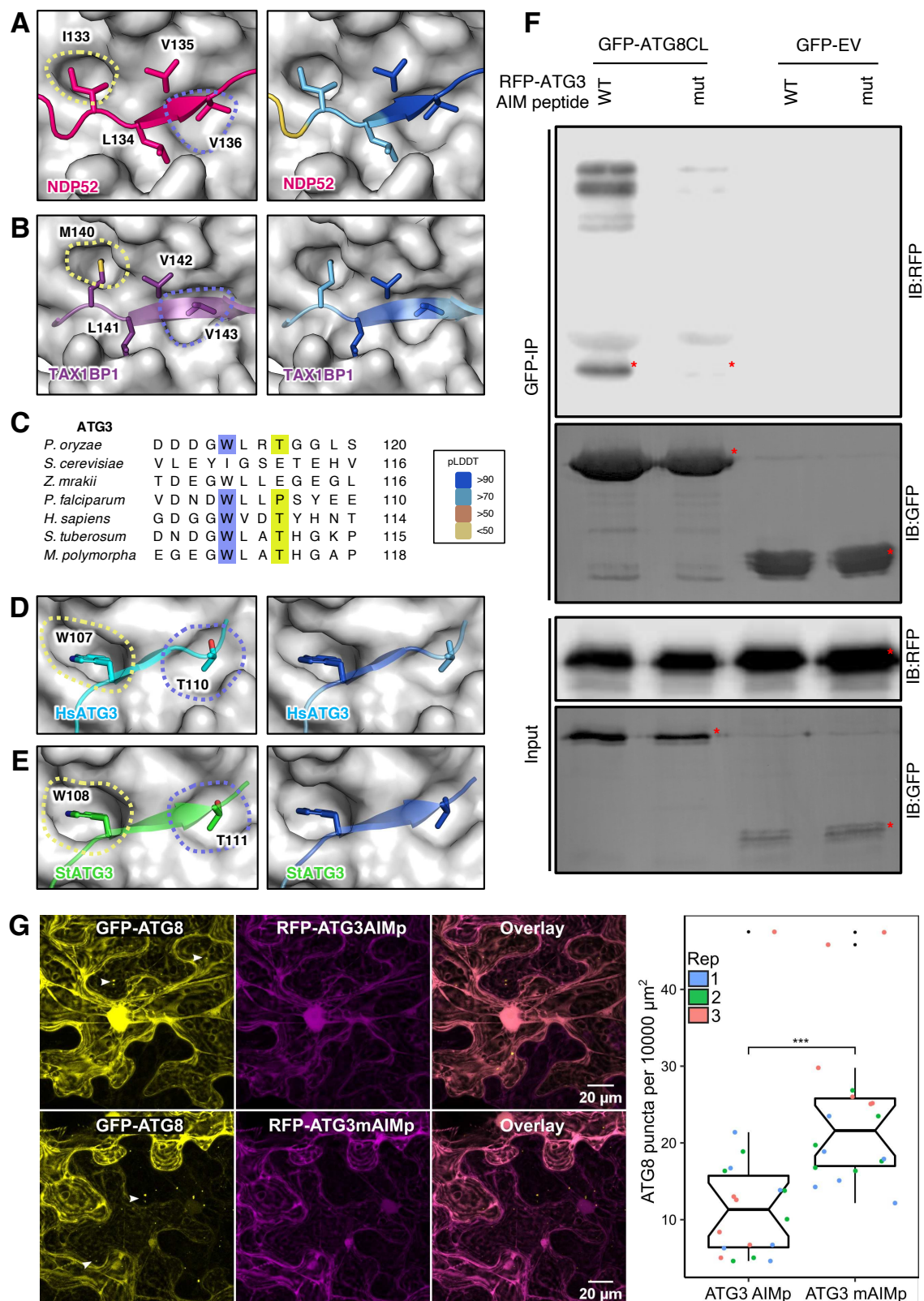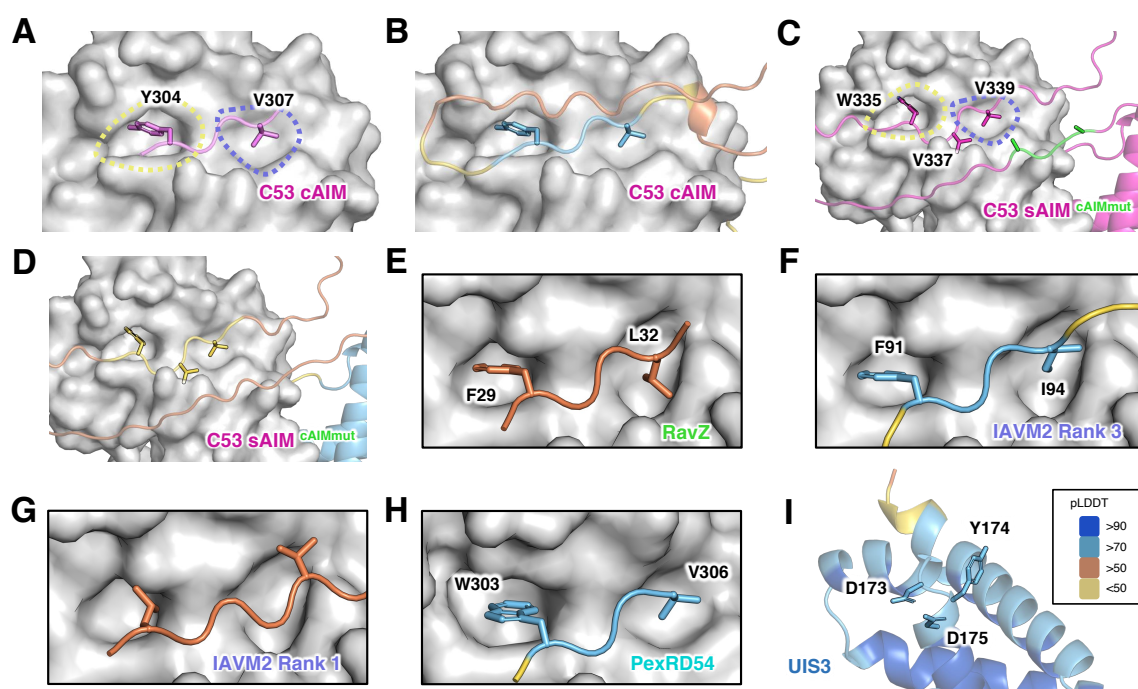
11

Figure 3. AF2-multimer can accurately predict experimentally validated non-canonical AIM/LIRs. (A) NDP52 LIR in complex with LC3C. Beyond the LIR pockets, W- (yellow) and L- (slate) sites, NDP52 (dark pink) utilises resides LEU134 and VAL135 to form additional interactions to provide specificity, as shown in literature. AF2-multimer predicts this complex with high confidence. **(B)** TAX1BP1 LIR in complex with LC3A. Beyond the LIR pockets W- (yellow) and L- (slate) sites, TAX1BP1 (purple) utilises resides LEU141 and VAL142 to form additional interactions to provide specificity, as shown in literature. AF2-multimer predicts this complex at a high confidence. **(C)** The non-canonical LIR of ATG3 is conserved in all but two species studied for this

12

project. **(D-E)** ATG3 (human – cyan, *S. tuberosum* – green) AIM/LIR in complex with ATG8-family members. Residues TRP and THR are arranged in a non-canonical AIM/LIR, as shown in literature. Right hand side of panels A, B, D and E are coloured based on the AF2-calculated prediction confidence score; predicted local distance difference test (pLDDT). Blue indicates regions of a protein with a score of over 90, meaning a very high confidence prediction. Scores between 70-90 are represented with light blue, which accounts for a high confidence score. Scores between 50-70 are considered low (orange) and anything below 50 is a very low confidence prediction (yellow). **(F)** In planta co-immunoprecipitation between ATG8CL and ATG3 AIM peptide and its mutant (WT/mut). GFP-ATG8CL or empty vector GFP control was transiently co-expressed with either RFP-wtAIMp or RFP-mutAIMp. Red asterisks indicate expected band sizes. **(G)** Confocal microscopy of wild-type *N. benthamiana* co-expressing GFP-ATG8CL and RFP-ATG3AIMpeptide (WT or mutant). White arrows indicate example puncta that was counted with different sizes. Images shown are maximal projections of 18 frames with 1 μm steps. Quantification of ATG8CL autophagosome puncta per 10000 μm^2 with WT and mutant (M) ATG3 AIM revealed a significantly higher number of puncta in sample co-expressed with mutated peptide compared to WT. Each repeat for the microscopy and quantification is represented by a different colour.

Another protein that was reported to have a non-canonical AIM/LIR is the Arabidopsis C53 (AtC53), which binds to ATG8 through several shuffled AIMs with the consensus sequence IDWG/D in addition to the canonical AIM YEIV [40, 41]. AF2-multimer analysis using AtC53 revealed the canonical AIM YEIV (Supplemental Figure S3A/B). To determine whether AF2-multimer can predict the non-canonical AIM in AtC53, we re-ran the analysis using an AtC53 sequence in which the canonical AIM had been mutated from YEIV to AEIA. In the absence of this canonical AIM, AF2-multimer was able to predict one of the three shuffled AIMs in C53 with the sequence WDVSV (amino acids 335–339) [40, 41]. According to the AF2-multimer model, while the W occupies the expected W-site (site one) on AtATG8a, the second V at position 339 occupies the L-site (site two), suggesting the existence of an extended non-canonical AIM with the sequence WDVSV (Supplemental Figure S3C/D).

Supplemental Figure 3. AF2-multimer determined confidence scores for C53 and cross-kingdom interactions with ATG8-family members. (A-B) Wild-type C53 preferentially interacts with ATG8A pockets via the canonical AIM (cAIM) formed by (304)YxxV(307) and this region of the protein within the complex is predicted at a high confidence. **(C)** Once the canonical AIM (cAIM - green) of C53 (pink) is mutated, a complex with ATG8A is formed by a shuffled AIM (sAIM), **(F)** however, this is not at high confidence, although it matches experimental observations. **(E)** RavZ, **(F)** IAVM2 model ranked third, with highest confidence at LIR region, **(G)** rank one model for IAVM2 and **(H)** PexRD54 complexes with respective ATG8-family members. **(I)** The suggested LIR of UIS3 (DYD) resides in an alpha-helix. All panels are coloured based on the AF2-calculated prediction confidence score; predicted local distance difference test (pLDDT). Blue indicates regions of a protein with a score of over 90, meaning a very high confidence prediction. Scores between 70-90 are represented with light blue, which accounts for a high confidence score. Scores between 50-70 are considered low (orange) and anything below 50 is a very low confidence prediction (yellow) (see Methods for further details).

Encouraged by the accurate prediction of canonical/non-canonical AIM/LIR motifs , we next checked if AF2-multimer could identify previously uncharacterized AIM//LIR residues. We decided to test ATG3 because, although it is well-established that it binds ATG8, how it does so remains controversial in different organisms [42]. While a canonical AIM is reported to mediate ATG8 binding in yeast, this AIM is conserved in fungi but not in human and plant ATG8s (Figure 3C). Intriguingly, AF2-multimer predicted a non-canonical AIM/LIR in both plant and human ATG3 proteins with WXXT sequence (Figure 3D/E). Consistent with the AF2 predictions, ATG3 from the malaria parasite (*Plasmodium falciparum)* was found to have a non-canonical AIM in this region with the WLLP sequence [43]. To validate the AF2-multimer prediction of the ATG3-ATG8 interaction, we cloned the AIM peptide region (X_{17}WLATX_4) of the potato ATG3 (StATG3) and its AIM mutant (AX_{17}ALAAX_4) and performed pulldown assays with the potato ATG8CL (StATG8CL) in plant cells. This revealed that wild type ATG3 AIM peptide binds ATG8CL whereas the mutated AIM peptide shows a significant reduction in ATG8CL interaction, validating the AF2-predicted model of non-canonical

14

AIM in plant ATG3 (Figure 3F). We have previously demonstrated that overproduction of functional AIM peptides blocks autophagosome formation in plants [44], which could be used as proxy to estimate AIM peptide ATG8 interaction *in planta*. Overexpression of the ATG3 AIM peptide reduced the formation of ATG8CL puncta (Figure 3G), indicating that it can block autophagy when overproduced in plant cells. In agreement with our findings a recent report showed human ATG3 interacts with LC3 through the non-canonical LIR sequence WXXT [45]. These results show that the AF2-multimer can predict both canonical/non-canonical AIM/LIRs and can serve as a powerful tool when combined with the molecular phylogeny analysis to provide new autophagy understanding.

AF2-multimer predicts cross-kingdom AIM/LIR-ATG8 interactions in host–pathogen interactions

Since many plants and human pathogens deploy virulence factors (also known as effectors) that target the host autophagy machinery, we investigated whether AF2-multimer can identify cross-kingdom ATG8–AIM/LIR associations in host–microbe interactions. To do so, we ran AF2-multimer using three pathogen virulence factors that bind host ATG8/LC3 proteins via experimentally validated AIM/LIR motifs. This included the RavZ protein secreted into human cells by the bacterial pathogen *Legionella pneumophila* [46]*,* Matrix protein 2 (M2) deployed by the Influenza A virus [47], and the PexRD54 protein secreted into plant cells by the Irish potato famine pathogen *Phytophthora infestans* [48]*.* For all three pathogen proteins carrying canonical AIM/LIRs, AF2-multimer successfully predicted the correct residues that mediate ATG8 interactions (Figure 4A-C, Supplemental Figure S3). AF2-multimer accurately predicted the functional C-terminal AIM (WEIV) in PexRD54 from among the two potential AIMs and the functional N-terminal LIR (FDLL) in RavZ from among 8 candidate LIRs, once again demonstrating the ability of AF2-multimer to determine the correct AIM from among multiple candidates. Of note, in the case of M2, the highest ranked AF2-model does not show a proper LIR (FVSI) fit in the LDS, whereas other models do show a perfect alignment of the core LIR residues, F and I, on the corresponding hydrophobic pockets on LC3. This is most likely because the confidence score of the first ranked model in the LIR/LDS region was much lower than that of other models (Supplemental Figure S3), indicating that confidence scores around the AIM/LIR region are more reliable than the overall protein model scores when determining AIM/LIR motifs by AF2-multimer. Additionally, we analysed the malaria (*Plasmodium berghei*) virulence factor PbUIS3, which does not carry a canonical LIR motif but instead binds to LC3 through an unidentified cryptic interface [49]. However, none of the AF2-multimer models showed any association of PbUIS3 residues with LC3-LDS, instead suggesting a different binding interface. According to these models, the previously identified PbUIS3 residues (D173, Y174, D175, E205, and K209) that are reported to be important for LC3 binding [49] are embedded in two separate

15

alpha-helices, which do not associate with LC3A (Supplemental Figure S3E), suggesting that mutations in these loci might indirectly limit LC3 binding. It would be interesting to investigate whether the AF2-multimer prediction of PbUIS3-LC3 association is correct or not. Despite limitations such as conditional ATG8 binding that can prevent AIM/LIR predictions, these analyses show that AF2-multimer can predict cross-kingdom AIM/LIR–ATG8 interactions.
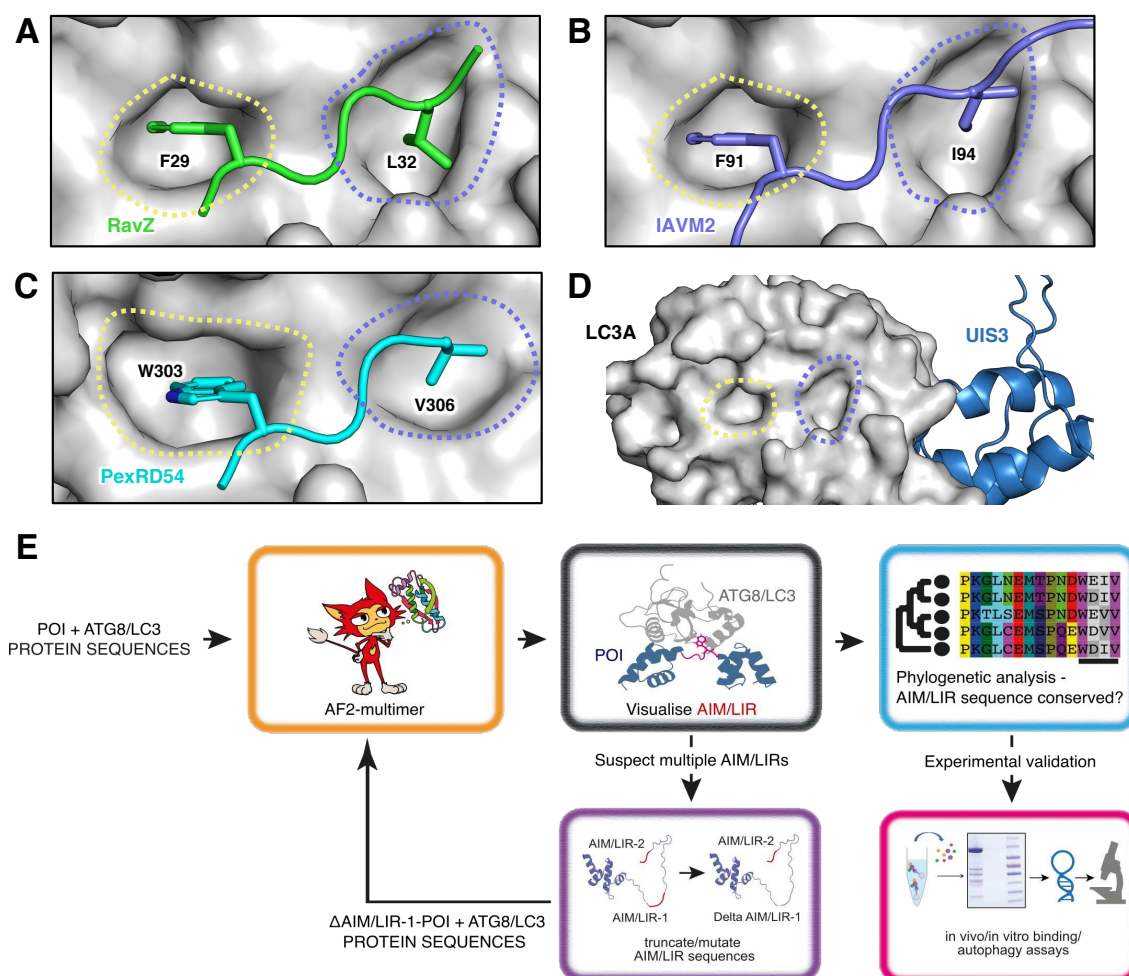

Figure 4. AF2-multimer guided discovery of cross-kingdom interactions. (A-C) RavZ (Legionella - green), M2 (Influenza – slate blue) and PexRD54 (*P. infestans* – cyan) interact with their respective ATG8-family member in humans (RavZ, M2) and plants (PexRD54) to produce a pathogenic phenotype. AF2-multimer correctly predicts interactions studied in literature. **(D)** UIS3 (Plasmodium – dark blue) does not interact with LC3A via its AIM/LIR pockets. **(E)** The pipeline for AF2-multimer and phylogeny guided discovery and validation of AIM/LIRs among various species. POI: protein of interest.

Discussion

We provide evidence that AF2-multimer can predict both canonical and non-canonical AIM/LIR motifs with high accuracy. In 36 proteins that carry at least one experimentally validated canonical AIM/LIR motif, AF2-multimer accurately identified 33 functional AIM/LIRs among 310 candidate

sequences that match the canonical AIM/LIR consensus (Figure 1 and Supplemental Table 1). AF2-multimer can also predict more than one functional AIM/LIR in a protein; however, it seems to prioritise displaying the N-terminal functional AIM/LIR. Nevertheless, this limitation can be circumvented by re-running AF2-multimer using truncated/mutated sequences of the first AIM/LIR (Figure 2). AF2-multimer accurately predicted all four experimentally validated non-canonical AIM/LIRs (Figure 3), an ability that is lacking in current AIM/LIR prediction algorithms. Remarkably, AF-2multimer accurately precited a previously uncharacterized non-canonical, functional AIM motif in plant ATG3, which we validated by ATG8 binding and autophagy assays (Figure 3). Furthermore, AF2-multimer performed well in identifying AIM/LIR motifs in 3 out of 4 tested pathogen virulence factors that target ATG8 members in their plant and human hosts, revealing that cross kingdom ATG8–LIR/AIM associations can also be predicted by AF2-multimer. AF2-multimer is therefore quite effective in identifying true AIM/LIR motifs among multiple other non-functional motifs in a protein of interest with high confidence, which minimises the experimental effort needed to validate these motifs. The confidence of these predictions can be strengthened by combining evolutionary information from phylogenetic analysis of protein sequences to determine AIM/LIR conservation, which should be considered when designing time-consuming validation experiments. The AF2-guided AIM/LIR prediction workflow consists of structural prediction of ATG8 and candidate proteins, phylogenetic analysis of predicted AIM/LIR motifs, and experimental validation by *in vitro*/*in vivo* binding assays (Figure 4E). If present, additional AIM/LIR residues can be identified by re-running the protocol using protein sequences that are mutated (*in silico*) in the primary AIM/LIR determined after the first run (Figure 4E).

A major advantage of AF2-multimer is that it provides spatial resolution of the AIM/LIR-ATG8 interaction, displaying not only structural alignment of the LIR/AIM residues on LDS/AIM docking site (ADS) but also additional associations through LIR/AIM flanking residues. This is important because it helps determine the extent to which AIM flanking residues condition AIM/LIR specificity towards certain ATG8 isoforms, thereby providing unprecedented insights into evolutionary studies focusing on autophagy regulation. This is especially important for discovering the mode of action of non-canonical AIM/LIR motifs, which are often overlooked as there are no known distinctive features to predict them by. Remarkably, using AF2-mitimer combined with experimental approaches and phylogenetic analysis, we were able to identify and validate a previously uncharacterized functional, non-canonical AIM motif in an IDPR of the plant ATG3 (Figure 3), indicating the AF2-multimer can inform new biology. As reported before [14], an emerging unifying pattern of the canonical and non-canonical AIMs is that they are typically located in IDPRs. Our AF2-multimer predictions are also in agreement with this view, as most AIM/LIR sequences that we analysed are in IDPRs (Figures 1-4).

According to our assessments, AF2-multimer correctly predicts the ATG8–AIM interface with greater confidence in some models than in others. However, a high-confidence structural prediction of a complete protein is not necessary to identify an ATG8 interface, as these regions can be ordered or disordered based on their interaction partners. The top-ranked AF2-multimer models correctly predicted the functional AIM/LIRs in all tested proteins except for the M2 protein from Influenza A virus. In the case of M2, the rank 1 model had a lower confidence score than the other models, which accurately showed the LIR–LC3 binding interface (Figure 4D). Therefore, models with the highest confidence scores in regions covering the AIM/LIR residues are more informative when determining AIM/LIR–ATG8 binding interfaces.

A current limitation of AF2-multimer is that it does not provide any information about the ATG8 binding affinity of the AIM/LIR motif. For instance, phosphorylation of residues flanking the LIR/AIM have been shown to improve ATG8 binding affinity [16, 17], but these cannot be determined by AF2-multimer. Nevertheless, since AF2-multimer was still able to identify AIM/LIRs without considering post-translational modifications, comparative analysis of the side chain interactions of flanking AIM/LIR residues revealed by AF2-multimer could provide further insights into AIM/LIR specificity towards different ATG8 members.

Despite its advantages, there is still room to improve AF2-multimer since not all AIM/LIR residues can be predicted. In three cases, i.e., the plant autophagy cargo receptor Joka2/NBR1 and the human Calreticulin and DVL2 proteins, AF2-multimer failed to predict the experimentally suggested AIM/LIRs [32, 44, 48]. We cannot rule out the possibility that ATG8/LC3 binding of such proteins is facilitated by post-translational modifications [50] and/or interactions with other proteins. Currently, such conditions are not considered by AF2-multimer. Intriguingly, the AF2-models suggest that other binding interfaces exist for these proteins. This raises the question of whether AF2-multimer can be used to predict additional, novel ATG8 binding interfaces. For instance, AF2-multimer predicted a previously unidentified LC3 binding interface in the malaria virulence factor PbUIS3, which lacks a typical LIR motif (Figure 4D). According to AF2-multimer models, the previously characterised PbUIS3 residues (D173, Y174, D175, E205, and K209) that are required for LC3 interaction are not located in the LC3 binding interface, indicating that they might indirectly affect LC3 binding. Consistent with this notion, these residues are also not conserved in other *Plasmodium* species [49], which contradicts the observations that AIM/LIR motifs are located in conserved, structurally disordered regions [14]. Thus, apart from identifying validated AIM/LIRs, AF2-multimer could be an important tool for determining new ATG8 binding interfaces, providing insights into cases where ATG8/LC3 interaction remains cryptic and developing new hypotheses that can be tested experimentally.

Here, we present an AF2-multimer–based framework for identifying AIM/LIR motifs that could be used to discover novel autophagy receptors, adaptors, or modulators as well as pathogen virulence factors that target ATG8 proteins. This framework, in turn, should help address key questions in plant and human autophagy, such as: How are specific cargoes captured and mobilised through selective autophagy? How does specific cargo selection help organisms withstand cellular and environmental stress? How are certain genetic diseases linked to defects in selective autophagy? What are the determinants of degradative versus secretory autophagy? And how do pathogens manipulate autophagy to promote diseases? The AI-guided quest for the discovery of autophagy modulators will substantially accelerate our understanding of the molecular basis of autophagy in all kingdoms of life.

Materials and Methods

Structural analysis using AF2

We analysed a total of 45 proteins, from 11 different species, that interacted with 12 different members of the ATG8-family of various species (Supplemental Table 1). Sequences were obtained from UniProt. Homologs of proteins studied were found using BLAST [51] and their sequences were aligned by Clustal Omega [52]. AF2-multimer [23] was used through a subscription to the Google Colab(https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/AlphaFold2.ipynb#scrollTo=svaADwocVdwl) following guidelines on the document [22]. Superposition of AlphaFold2 predictions on known structures was performed using the align command in PyMOL (The PyMOL Molecular Graphics System, Version 2.3.5 Schrödinger, LLC). Where indicated, predictions of AIM/LIR interactions are coloured according to the AlphaFold2 produced per-residue confidence metric called the local distance difference test (pLDDT), which corresponds to the model's predicted score on the lDDT-Cα metric [53]. The scale ranges from 0-100, where 100 corresponds to values of highest confidence. Models were coloured by this score in PyMOL using a script generated by J. Murray (Imperial College London).

Plasmid constructs

RFP- and GFP- ATG8CL used in this study were published previously [44]. C-terminal GFP-tagged ATI2 construct was generated by Gibson assembly of PCR fragment amplified from *S. tuberosum* cDNA into EcoRV linearized pKGC3S GFP expression vector. The DNA fragment for ATI2 AIM1 peptide was custom synthesised (GeneWiz) and inserted into the pKGN3S vector to generate an N-

19

terminal GFP expression construct using Gibson assembly. ATI2-AIM1_(17)AEVA(20), ATI2-AIM2_(265)AQVA(268) and ATIAIM1+2 mutant fragments were generated using site directed mutagenesis (SDM) polymerase chain reaction (PCR) amplification from wild-type construct. Templates were then eliminated by one-hour Dpn-I (New England Biolabs) restriction digestion at 37°C and the PCR products of mutants were inserted into EcoRV digested C-terminal GFP expression vectors using Gibson assembly. ATG3 AIM peptide DNA fragments were custom synthesised (GeneWiz) and inserted into the pKRN3S vector to generate an N-terminal RFP expression construct using Gibson assembly. Expression vectors were transformed via electroporation into Agrobacterium strain GV3101 for plant expression.

Plant material and growth conditions

N. benthamiana plants were grown and maintained in a greenhouse with high light intensity (16 hr light/8 hr dark photoperiod) at 22–24°C.

Co-immunoprecipitation experiments and immunoblot analysis

Proteins were transiently expressed by agroinfiltration *N. benthamiana* leaves and harvested 2-days post agroinfiltration. Protein extraction, purification and western blot analysis steps were performed as described previously [54]. In brief, 2g of leaf tissue was grinded in 4mL extraction buffer (25 mM Tris-HCl, pH 7.5, 1 mM EDTA, 150 mM NaCl, 10% glycerol (v/v), and 10 mM DTT, 0.1% IGEPAL in the presence of plant protease inhibitor cocktail (Sigma-Aldrich) and 2% polyvinylpolypyrrolidone). Following centrifugation at 12000 RPM for two-times 20 minutes (filtration in between), the resultant supernatant was incubated with RFP beads (Chromotek) for 1 hour at 4°C. Beads were washed at 800g three times prior to elution with an elution buffer (4xLaemmli Buffer (BioRad) and DTT). Proteins were eluted at 70°C for 5 minutes. After gel electrophoresis and transfer to PVDF membrane, polyclonal anti-GFP (Chromotek) produced in rabbit, monoclonal anti-RFP (Chromotek) produced in mouse and monoclonal anti-GFP (Chromotek) produced in rat were used as primary antibodies. For secondary antibodies anti-mouse antibody (Sigma-Aldrich), anti-rabbit (Sigma-Aldrich) and anti-rat (Sigma-Aldrich) antibodies were used.

Confocal microscopy and quantitative analysis of GFP-ATG8CL autophagosome puncta

Imaging was performed using Leica Stellaris 5 inverted confocal microscope (Leica Microsystems) using 63x water immersion objective. All microscopy analyses were carried out on live leaf tissue 3 days after agroinfiltration. Leaf discs of *N. benthamiana* were cut and mounted onto Carolina observation gel (Carolina Biological Supply Company) to minimise the damage. Specific excitation wavelengths and filters for emission spectra were set as described previously [55]. GFP and RFP probes were excited using 488 and 561 nm laser diodes and their fluorescent emissions detected at 495–550 and 570–620 nm, respectively. Maximum intensity projections of Z-stack images were presented in each figure. Image analysis was performed using Fiji and Inkscape. For quantifying GFP-ATG8 puncta, maximum-intensity projection of each confocal microscopy image was generated using Fiji [56]. The total cell surface area was measured after removing, if there was any, the stomata or pavement cells from the images using the "Freehand selections" tool of Fiji, if the cells were not expressing GFP-ATG8. Since the signal from the ATG8 puncta has been higher than the cytoplasmic-ATG8, "Find maxima" tool was used to measure the number of puncta, with the "prominence" setting between 45 and 60 depending on the signal intensity of each image. The relative number of puncta per 10000 μm2 was then calculated to plot a graph. The statistical significance of any differences in means was found using the Wilcoxon test in R, because the Shapiro–Wilk test showed the data did not follow normal distribution. Three asterisks (***) in the box plot indicates that the p value is smaller than 0.001.

Acknowledgements

References

1.   Yang Z, Klionsky DJ. An overview of the molecular mechanism of autophagy. Curr Top Microbiol Immunol. 2009;335:1-32.

2.   Suzuki K, Kubota Y, Sekito T, Ohsumi Y. Hierarchy of Atg proteins in pre-autophagosomal structure organization. Genes Cells. 2007;12(2):209-18.

3.   Parzych KR, Klionsky DJ. An overview of autophagy: morphology, mechanism, and regulation. Antioxid Redox Signal. 2014;20(3):460-73.

4.   New J, Thomas SM. Autophagy-dependent secretion: mechanism, factors secreted, and disease implications. Autophagy. 2019;15(10):1682-93.

5.   Shpilka T, Weidberg H, Pietrokovski S, Elazar Z. Atg8: an autophagy-related ubiquitin-like protein family. Genome Biol. 2011;12(7):226.

6.   Kriegenburg F, Ungermann C, Reggiori F. Coordination of Autophagosome-Lysosome Fusion by Atg8 Family Members. Curr Biol. 2018;28(8):R512-R8.

7.   Kellner R, De la Concepcion JC, Maqbool A, Kamoun S, Dagdas YF. ATG8 Expansion: A Driver of Selective Autophagy Diversification? Trends Plant Sci. 2017;22(3):204-14.

8.   Taherbhoy AM, Tait SW, Kaiser SE, Williams AH, Deng A, Nourse A, et al. Atg8 transfer from Atg7 to Atg3: a distinctive E1-E2 architecture and mechanism in the autophagy pathway. Mol Cell. 2011;44(3):451-61.

9.   Nakatogawa H, Ohbayashi S, Sakoh-Nakatogawa M, Kakuta S, Suzuki SW, Kirisako H, et al. The autophagy-related protein kinase Atg1 interacts with the ubiquitin-like protein Atg8 via the Atg8 family interacting motif to facilitate autophagosome formation. J Biol Chem. 2012;287(34):28503-7.

10.   Fracchiolla D, Sawa-Makarska J, Zens B, Ruiter A, Zaffagnini G, Brezovich A, et al. Mechanism of cargo-directed Atg8 conjugation during selective autophagy. Elife. 2016;5.

11.   Stolz A, Ernst A, Dikic I. Cargo recognition and trafficking in selective autophagy. Nat Cell Biol. 2014;16(6):495-501.

12.   Pankiv S, Alemu EA, Brech A, Bruun JA, Lamark T, Overvatn A, et al. FYCO1 is a Rab7 effector that binds to LC3 and PI3P to mediate microtubule plus end-directed vesicle transport. J Cell Biol. 2010;188(2):253-69.

13.   McEwan DG, Popovic D, Gubas A, Terawaki S, Suzuki H, Stadel D, et al. PLEKHM1 regulates autophagosome-lysosome fusion through HOPS complex and LC3/GABARAP proteins. Mol Cell. 2015;57(1):39-54.

14.   Popelka H, Klionsky DJ. Analysis of the native conformation of the LIR/AIM motif in the Atg8/LC3/GABARAP-binding proteins. Autophagy. 2015;11(12):2153-9.

15.   Fracchiolla D, Sawa-Makarska J, Martens S. Beyond Atg8 binding: The role of AIM/LIR motifs in autophagy. Autophagy. 2017;13(5):978-9.

16.   Wirth M, Mouilleron S, Zhang W, Sjottem E, Princely Abudu Y, Jain A, et al. Phosphorylation of the LIR Domain of SCOC Modulates ATG8 Binding Affinity and Specificity. J Mol Biol. 2021;433(13):166987.

17.   Katsuragi Y, Ichimura Y, Komatsu M. p62/SQSTM1 functions as a signaling hub and an autophagy adaptor. FEBS J. 2015;282(24):4672-8.

18.   von Muhlinen N, Akutsu M, Ravenhill BJ, Foeglein A, Bloor S, Rutherford TJ, et al. LC3C, bound selectively by a noncanonical LIR motif in NDP52, is required for antibacterial autophagy. Mol Cell. 2012;48(3):329-42.

19.   Kalvari I, Tsompanis S, Mulakkal NC, Osgood R, Johansen T, Nezis IP, et al. iLIR: A web resource for prediction of Atg8-family interacting proteins. Autophagy. 2014;10(5):913-25.

20.   Xie Q, Tzfadia O, Levy M, Weithorn E, Peled-Zehavi H, Van Parys T, et al. hfAIM: A reliable bioinformatics approach for in silico genome-wide identification of autophagy-associated Atg8-interacting motifs in various organisms. Autophagy. 2016;12(5):876-87.

21.   Han Z, Zhang W, Ning W, Wang C, Deng W, Li Z, et al. Model-based analysis uncovers mutations altering autophagy selectivity in human cancer. Nat Commun. 2021;12(1):3258.

22.   Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. Nature. 2021;596(7873):583-9.

23.   Evans R, O'Neill M, Pritzel A, Antropova N, Senior A, Green T, et al. Protein complex prediction with AlphaFold-Multimer. bioRxiv. 2022:2021.10.04.463034.

24.   Maqbool A, Hughes RK, Dagdas YF, Tregidgo N, Zess E, Belhaj K, et al. Structural Basis of Host Autophagy-related Protein 8 (ATG8) Binding by the Irish Potato Famine Pathogen Effector Protein PexRD54. J Biol Chem. 2016;291(38):20270-82.

25.   Knight D, Harris R, McAlister MS, Phelan JP, Geddes S, Moss SJ, et al. The X-ray crystal structure and putative ligand-derived peptide binding properties of gamma-aminobutyric acid receptor type A receptor-associated protein. J Biol Chem. 2002;277(7):5556-61.

26.   Wilfling F, Lee CW, Erdmann PS, Zheng Y, Sherpa D, Jentsch S, et al. A Selective Autophagy Pathway for Phase-Separated Endocytic Protein Deposits. Mol Cell. 2020;80(5):764-78 e7.

27.   Zientara-Rytter K, Sirko A. Selective autophagy receptor Joka2 co-localizes with cytoskeleton in plant cells. Plant Signal Behav. 2014;9(3):e28523.

28.   Thielmann Y, Weiergraber OH, Mohrluder J, Willbold D. Structural framework of the GABARAP-calreticulin interface--implications for substrate binding to endoplasmic reticulum chaperones. FEBS J. 2009;276(4):1140-52.

29.   Chouquet A, Paidassi H, Ling WL, Frachet P, Houen G, Arlaud GJ, et al. X-ray structure of the human calreticulin globular domain reveals a peptide-binding area and suggests a multi-molecular mechanism. PLoS One. 2011;6(3):e17886.

30.   Wesch N, Kirkin V, Rogov VV. Atg8-Family Proteins-Structural Features and Molecular Interactions in Autophagy and Beyond. Cells. 2020;9(9).

31.   Wirth M, Zhang W, Razi M, Nyoni L, Joshi D, O'Reilly N, et al. Molecular determinants regulating selective binding of autophagy adapters and receptors to ATG8 proteins. Nat Commun. 2019;10(1):2055.

32.   Gao C, Cao W, Bao L, Zuo W, Xie G, Cai T, et al. Autophagy negatively regulates Wnt signalling by promoting Dishevelled degradation. Nat Cell Biol. 2010;12(8):781-90.

33.   Sun A, Wei J, Childress C, Shaw JHt, Peng K, Shao G, et al. The E3 ubiquitin ligase NEDD4 is an LC3-interactive protein and regulates autophagy. Autophagy. 2017;13(3):522-37.

34.   Qiu Y, Zheng Y, Wu KP, Schulman BA. Insights into links between autophagy and the ubiquitin system from the structure of LC3B bound to the LIR motif from the E3 ligase NEDD4. Protein Sci. 2017;26(8):1674-80.

35.   Michaeli S, Honig A, Levanony H, Peled-Zehavi H, Galili G. Arabidopsis ATG8-INTERACTING PROTEIN1 is involved in autophagy-dependent vesicular trafficking of plastid proteins to the vacuole. Plant Cell. 2014;26(10):4084-101.

36.   Sjogaard IMZ, Bressendorff S, Prestel A, Kausika S, Oksbjerg E, Kragelund BB, et al. The transmembrane autophagy cargo receptors ATI1 and ATI2 interact with ATG8 through intrinsically disordered regions with distinct biophysical properties. Biochem J. 2019;476(3):449-65.

37.   Chatzichristofi A, Sagris V, Pallaris A, Eftychiou M, Kalvari I, Price N, et al. LIRcentral: a manually curated online database of experimentally validated functional LIR-motifs. bioRxiv. 2022:2022.06.21.496832.

38.   Tumbarello DA, Manna PT, Allen M, Bycroft M, Arden SD, Kendrick-Jones J, et al. Correction: The Autophagy Receptor TAX1BP1 and the Molecular Motor Myosin VI Are Required for Clearance of Salmonella Typhimurium by Autophagy. PLoS Pathog. 2016;12(1):e1005433.

39.   Mochida K, Yamasaki A, Matoba K, Kirisako H, Noda NN, Nakatogawa H. Super-assembly of ER-phagy receptor Atg40 induces local ER remodeling at contacts with forming autophagosomal membranes. Nat Commun. 2020;11(1):3306.

40.   Stephani M, Picchianti L, Gajic A, Beveridge R, Skarwan E, Sanchez de Medina Hernandez V, et al. A cross-kingdom conserved ER-phagy receptor maintains endoplasmic reticulum homeostasis during stress. Elife. 2020;9.

41.   Picchianti L, de Medina Hernández VS, Zhan N, Irwin NAT, Stephani M, Hornegger H, et al. Shuffled ATG8 interacting motifs form an ancestral bridge between UFMylation and C53-mediated autophagy. bioRxiv. 2022:2022.04.26.489478.

42.   Fang D, Xie H, Hu T, Shan H, Li M. Binding Features and Functions of ATG3. Front Cell Dev Biol. 2021;9:685625.

43.   Hain AU, Weltzer RR, Hammond H, Jayabalasingham B, Dinglasan RR, Graham DR, et al. Structural characterization and inhibition of the Plasmodium Atg8-Atg3 interaction. J Struct Biol. 2012;180(3):551-62.

44.   Pandey P, Leary AY, Tumtas Y, Savage Z, Dagvadorj B, Duggan C, et al. An oomycete effector subverts host vesicle trafficking to channel starvation-induced autophagy to the pathogen interface. Elife. 2021;10.

45.   Farnung J, Muhar M, Liang JR, Tolmachova KA, Benoit RM, Corn JE, et al. Human ATG3 contains a non-canonical LIR motif crucial for its enzymatic activity in autophagy. bioRxiv. 2022:2022.08.02.502437.

46.   Park SW, Jun YW, Jeon P, Lee YK, Park JH, Lee SH, et al. LIR motifs and the membrane-targeting domain are complementary in the function of RavZ. BMB Rep. 2019;52(12):700-5.

47.   Beale R, Wise H, Stuart A, Ravenhill BJ, Digard P, Randow F. A LC3-interacting motif in the influenza A virus M2 protein is required to subvert autophagy and maintain virion stability. Cell Host Microbe. 2014;15(2):239-47.

48.   Dagdas YF, Belhaj K, Maqbool A, Chaparro-Garcia A, Pandey P, Petre B, et al. An effector of the Irish potato famine pathogen antagonizes a host autophagy cargo receptor. Elife. 2016;5.

49.   Real E, Rodrigues L, Cabal GG, Enguita FJ, Mancio-Silva L, Mello-Vieira J, et al. Plasmodium UIS3 sequesters host LC3 to avoid elimination by autophagy in hepatocytes. Nat Microbiol. 2018;3(1):17-25.

50.   Xie Y, Kang R, Sun X, Zhong M, Huang J, Klionsky DJ, et al. Posttranslational modification of autophagy-related proteins in macroautophagy. Autophagy. 2015;11(1):28-45.

51.   Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990;215(3):403-10.

52.   Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol. 2011;7:539.

53.   Mariani V, Biasini M, Barbato A, Schwede T. lDDT: a local superposition-free score for comparing protein structures and models using distance difference tests. Bioinformatics. 2013;29(21):2722-8.

54.   Bozkurt TO, Schornack S, Win J, Shindo T, Ilyas M, Oliva R, et al. Phytophthora infestans effector AVRblb2 prevents secretion of a plant immune protease at the haustorial interface. Proc Natl Acad Sci U S A. 2011;108(51):20832-7.

55.   Koh S, Andre A, Edwards H, Ehrhardt D, Somerville S. Arabidopsis thaliana subcellular responses to compatible Erysiphe cichoracearum infections. Plant J. 2005;44(3):516-29.

56.   Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, et al. Fiji: an open-source platform for biological-image analysis. Nat Methods. 2012;9(7):676-82.