nature REVIEWS MOLECULAR CELL BIOLOGY

Ubiquitin and ubiquitin-like proteins

Rebecca L. Welchman, Tim Soane, Paul W. Sheppard and R. John Mayer

Protein ubiquitylation is a recognized signal for protein degradation. However, it is increasingly realized that ubiquitin conjugation to proteins can be used for many other purposes, and there are many ubiquitin-like proteins that control the activities of proteins. The central structural element of these post-translational modifications is the ubiquitin superfold and, as well as being small conjugatable protein modifiers, ubiquitin superfolds can be domains that are genetically built into much larger proteins. An encompassing term for each of these structural folds is 'ubiquiton'. Ubiquitons have various functions, most of which are unrelated to protein degradation, and some ubiquitons have little homology to ubiquitin. In the future, we expect to see that numerous regulatory proteins are conjugated to ubiquitons to enhance the specificity of protein interactions.

Ubiquitin-like protein nomenclature

Ubls	Alternative protein names	Swiss-Prot accession number	Number of amino acids*
Ubiquitin	–	P62988	76
NEDD8	Neddylin, Rub1	Q15843	81
SUMO-1	Smt3c, PIC1, UBL1, GMP1, Sentrin	P63165	101
SUMO-2	Smt3a	P55854	103
SUMO-3	Smt3b, Sentrin-2, HSMT3	P61956	95
ISG15	UCRP	P05161	164
FAT10	Ubiquitin D, diubiquitin	O15205	165
FUB1	FUBI	P35544	74
UBL5	Hub1	Q9BZL1	73
URM1	C9ORF74	Q9BTM9	101
ATG8	MAP-LC3, γ-aminobutyric-acid-receptor-associated protein	O95166	117
ATG12	APG12	O94817	140

*The number of amino acids refers to the processed, mature proteins.

Mammalian Ubls and their conjugation-cascade components

Ubls	E1	E2	E3	DUB*
Ubiquitin	UBE1	Many‡	Many‡	Many‡
NEDD8	UBE1c, ULA1	UBCH12	?	SENP8, CSN5
SUMO-1, -2, -3	ULE1a, ULA1b	UBCH9 homologue	PIAS, Pc2, RanBP2	SENP1, 2, 3, 6
ISG15	UBE1L	UBC8	?	USP18
FAT10	?	?	?	?
FUB1	?	?	?	?
UBL5	?	?	?	?
URM1	UBA4	?	?	?
ATG8	ATG7	ATG3	?	ATG4
ATG12	ATG7	ATG10	?	?

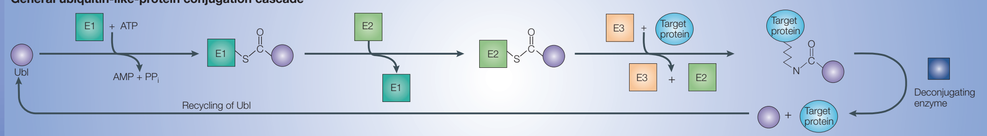
*DUB or corresponding deconjugating enzyme. ‡In mammals, there might be 50–70 E2s, >500 E3s and 50–70 DUBs (the exact numbers are still being determined). Question marks highlight components that are unknown to date.

Phylogenetic conservation of ubiquitin-like proteins and their activators*

	Homo sapiens	*Saccharomyces cerevisiae*	*Ashbya gossypii*	*Cyanidioschyzon merolae*	*Encephalitozoon cuniculi*
Ubls					
Ubiquitin	Several copies	Several copies	Several copies	Several copies	Several copies
NEDD8	1	1	1	1	1
SUMO	4	1	1	1	1
URM1	1	1	1	1	1
ATG12	1	1	1	0	0
Ubl activators					
UBE1	+	+	+	+	+
UBE1c, ULA1	+, +	+, +	+, +	+, +	s
ULE1a, ULA1b	+, +	+, +	+, +	+, +	s
UBA4	+	+	+	+	+
ATG7	+	+	+	–	–
Other activators	3	2	1	0	0

*This table shows the number of genes there are for the various Ubls in each organism, as well as for other Ubl activators (last row). *H. sapiens* has ~23,000 genes in total; *S. cerevisiae* has 6,000; the yeast *A. gossypii* (which is closely related to *S. cerevisiae*, but without genome duplication) has 3,900 non-duplicated genes; the green alga *C. merolae* has 4,700; and the microsporian *E. cuniculi* has 1,800. For the Ubl activators, the '+' symbol means present in the genome, '–' means absent from the genome, and 's' means shared (for example, *E. cuniculi* has only one gene for ULE1a and ULA1b, which is probably used for both enzyme functions). This information was kindly provided by Kay Hofmann (MEMOREC, Cologne, Germany).

General ubiquitin-like-protein conjugation cascade*



*The E3 usually facilitates direct Ubl transfer from the E2 to the target protein without forming a thioester bond itself, although, in a minority of cases, E3s directly form thioesters with the Ubl before the Ubl is conjugated to the target protein. In addition, transfer sometimes occurs without the involvement of an E3.

Percentage sequence identity between human ubiquitin-like proteins

	Ubiquitin	NEDD8	SUMO-1	SUMO-2	SUMO-3	ISG15a*	ISG15b*	FAT10a*	FAT10b*	FUB1	UBL5	URM1	ATG8	ATG12
Ubiquitin	100	58‡	18	16	16	29‡	37‡	29‡	36‡	37‡	22‡	12	10	17
NEDD8		100	21‡	18	18	28‡	26‡	26‡	29‡	24‡	18	13	6	17
SUMO-1			100	54‡	54‡	16	20‡	16	15	23‡	13	10	13	16
SUMO-2				100	100‡	16	13	8	11	19	15	11	13	10
SUMO-3					100	16	13	8	11	19	15	11	13	10
ISG15a*						100	26‡	20‡	22‡	18	17	11	8	9
ISG15b*							100	21‡	23‡	35‡	14	17	10	11
FAT10a*								100	20‡	27‡	20‡	11	12	8
FAT10b*									100	31‡	19	13	9	9
FUB1										100	14	10	11	10
UBL5											100	9	11	10
URM1												100	9	3
ATG8													100	26‡
ATG12														100

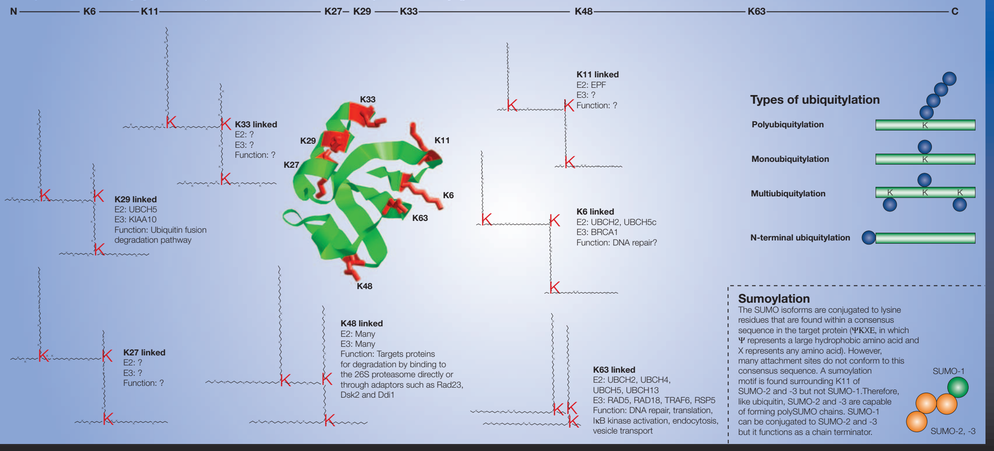
*The non-ubiquitin-like parts of the proteins were removed before the analysis: ISG15 and FAT10 are therefore shown twice, because they contain two ubiquitin-like domains each. ‡These values do not strictly correlate with identity, but are based on a statistical calculation (p<0.01). These ubiquitin alignments were calculated using the Needleman and Wunsch method in 'global' mode, and this information was kindly provided by Kay Hofmann (MEMOREC, Cologne, Germany).

Percentage sequence identity between yeast ubiquitin-like proteins

	Ubiquitin	Rub1	Smt3	Hub1	Urm1	Atg8	Atg12	Human*
Ubiquitin	100							96‡
Rub1	53‡	100						59‡
Smt3	16	18	100					51‡
Hub1	21‡	22‡	11	100				64‡
Urm1	19	21	22	15	100			42‡
Atg8	18	21	15	20	20	100		58‡
Atg12	15	9	17	14	16	21‡	100	28‡

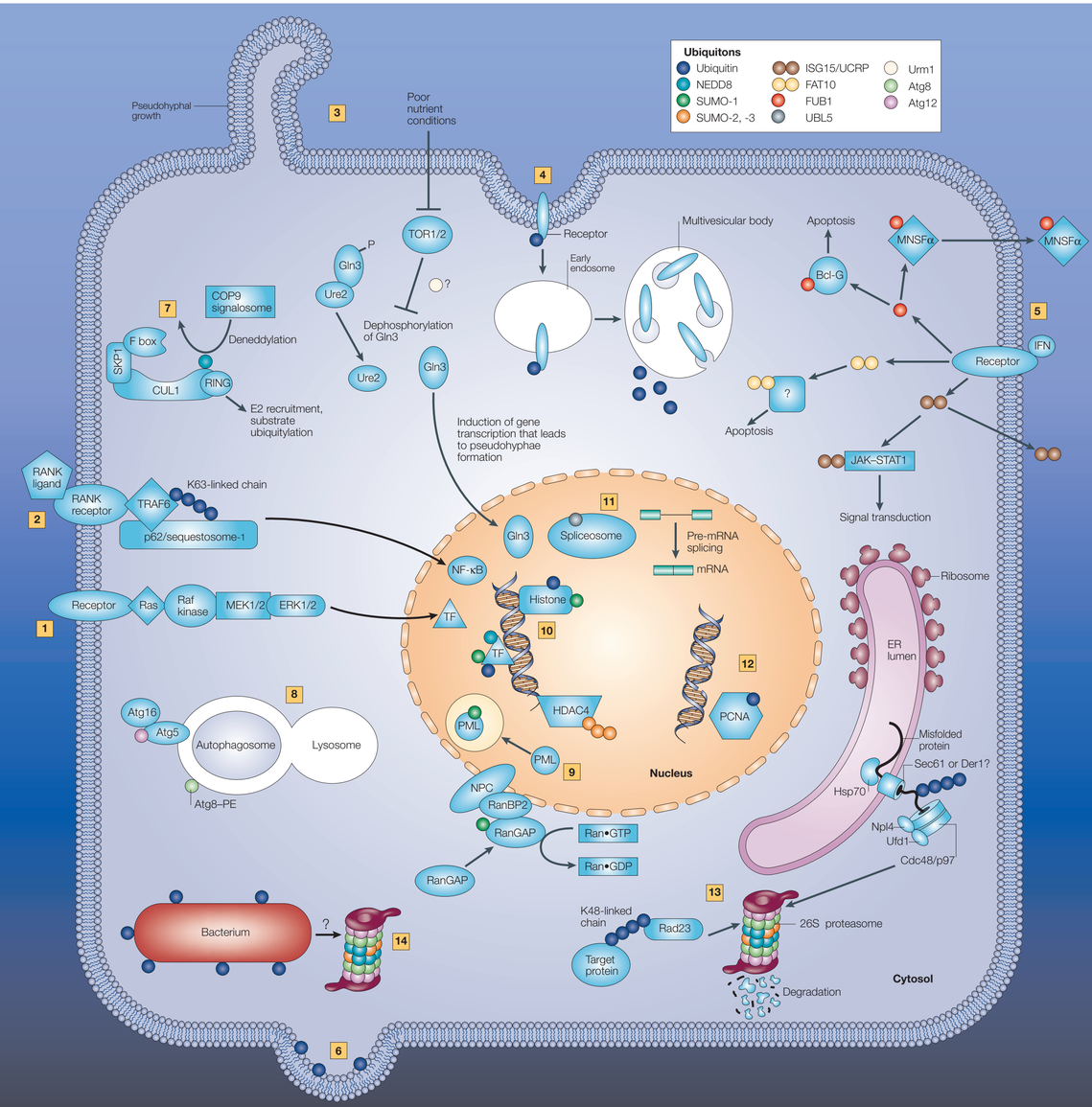
*The final column shows the homology between yeast ubiquitin/Ubls and human ubiquitin. ‡These values do not strictly correlate with sequence identity, but are based on a statistical calculation (p<0.01). These ubiquitin alignments were calculated using the Needleman and Wunsch method in 'global' mode, and this information was kindly provided by Kay Hofmann (MEMOREC, Cologne, Germany).

Ubiquitin has seven lysyl residues, all of which can function as potential conjugation sites



Sumoylation

The SUMO isoforms are conjugated to lysine residues that are found within a consensus sequence in the target protein (ΨKXE, in which Ψ represents a large hydrophobic amino acid and X represents any amino acid). However, many attachment sites do not conform to this consensus sequence. A sumoylation motif is found surrounding K11 of SUMO-2 and -3 but not SUMO-1.Therefore, like ubiquitin, SUMO-2 and -3 are capable of forming polySUMO chains. SUMO-1 can be conjugated to SUMO-2 and -3 but it functions as a chain terminator.



Ubiquitons

Example functions of ubiquitons:

1 Genetically built-in ubiquitons and signalling. Raf kinase contains a built-in ubiquiton in its Ras-binding domain, which facilitates protein–protein interactions.

2 K63-linked ubiquitin chains and signalling. The adaptor protein p62/sequestosome-1 binds K63-linked-chain-modified TRAF6.

3 Urm1 and nutrient sensing. In *Saccharomyces cerevisiae*, Urm1 is involved in regulating nutrient sensing and pseudohyphal growth through the TOR signalling pathway. TOR maintains Gln3 in a phosphorylated state, so that it remains bound to Ure2. When TOR is inactivated, Gln3 is dephosphorylated and released.

4 Ubiquitin and endocytosis. Receptor monoubiquitylation induces endocytosis and the generation of MVBs. Some receptors are multiubiquitylated or polyubiquitylated with short K63-linked chains.

5 ISG15/UCRP, FAT10, FUB1 and the immune system. These ubiquitons are expressed in higher eukaryotes with immune systems, and are induced by interferons. They are involved in downstream signalling, and FUB1 and ISG15 are also secreted as lymphokines.

6 Ubiquitin and phospholipids. Retroviruses subvert the MVB system to exit cells, and the release of ubiquitin from phosphatidyl–ubiquitin might facilitate ubiquitylation at membranes.

7 NEDD8 and SCF E3 ligases. NEDD8-modified CUL1 cannot associate with an SCF inhibitor. The COP9 signalosome deneddylates CUL1, which allows inhibitor binding and SKP1–F-box displacement.

8 Atg12, Atg8 and autophagy. Atg12–Atg5 and Atg8–PE localize to forming autophagosomal membranes. Atg16 binds Atg12–Atg5 to trigger autophagosome formation.

9 SUMO, PML bodies and nuclear pores. SUMO-1-modified PML localizes to nuclear PML bodies, whereas SUMO-1-modified RanGAP localizes to NPCs. SUMO-2 and -3 modify many of the same proteins as SUMO-1, including HDAC4. Unlike SUMO-1, SUMO-2 and -3 can form chains (function unknown).

10 Ubiquitons and transcription. Chromatin remodelling and transcription are regulated by ubiquitons such as ubiquitin, SUMO and NEDD8.

11 UBL5 and pre-mRNA splicing. UBL5 (Hub1 in *S. cerevisiae*) co-immunoprecipitates with spliceosome-complex intermediates (the linkage might not be covalent).

12 Ubiquitons and DNA repair. The DNA-replication processivity factor PCNA can be ubiquitylated, polyubiquitylated or sumoylated to induce different repair processes.

13 Ubiquitin–proteasome-mediated protein degradation. Polyubiquitylated cytosolic proteins can directly bind to the 26S proteasome or be delivered by adaptors (for example, Rad23). In ERAD, misfolded ER proteins are retrotranslocated (potentially through the Sec61 complex or putative Der1 channel), ubiquitylated, bound by Cdc48/p97–Ufd1–Npl4 and delivered to the proteasome. Interestingly, Rad23 contains a built-in ubiquiton.

14 Ubiquitin and intracellular bacteria. Intracellular bacteria might use the ubiquitin–proteasome system to enter the cytosol from intracellular vesicles, or this system might be involved in bacterial elimination.

For further information, please refer to the Review by Rebecca L. Welchman, Colin Gordon and R. John Mayer in the August 2005 issue of *Nature Reviews Molecular Cell Biology*.

Apg/APG/Atg/ATG, autophagy; Bac5, Bactenecin 5; Bcl-G, B-cell lymphoma-G; BRCA1, breast-cancer-susceptibility protein-1; C9ORF74, chromosome 9 open reading frame 74; CSN5, COP9 signalosome complex subunit 5; CUE, coupling of ubiquitin conjugation to ER degradation; CUL, cullin; Ddi1, DNA damage molecule-1; Der1, Derlin-1; Dsk2, dominant suppressor of Kar2; DUB, deubiquitylating enzyme; E1, enzyme-1 (ubiquitin-activating enzyme); E2, enzyme-2 (ubiquitin-conjugating enzyme); E3, enzyme-3 (ubiquitin-protein ligase); EPF, endemic pemphigus foliaceus; ER, endoplasmic reticulum; ERAD, ER-associated degradation; ERK, extracellular signal-regulated kinase; FUB1, Fau ubiquitin-like protein-1; GMP1, GAP-modifying protein-1; HDAC4, histone deacetylase-4; Hsp70, 70-kDa heat-shock protein; Hub1, homologous to ubiquitin-1; IκB, inhibitor of NF-κB; IFN, interferon; ISG15, interferon-stimulated gene-15; JAK, Janus kinase; MAP-LC3, microtubule-associated protein light chain 3; MEK, mitogen-activated protein kinase (MAPK) and extracellular signal-regulated kinase (ERK) kinase; MNSF, monoclonal nonspecific suppressor factor; MVB, multivesicular body; NEDD8, neuronal-precursor-cell-expressed developmentally downregulated protein-8; NF-κB, nuclear factor-κB; NPC, nuclear pore complex; P, phosphate; Pc2, Polycomb protein-2; PCNA, proliferating cell nuclear antigen; PE, phosphatidylethanolamine; PIAS, protein inhibitor of activated STAT; PIC1, PML-interacting clone-1; PML, promyelocytic leukaemia protein; PP_i, pyrophosphate; Rad/RAD, radiation gene; RanBP2, Ran-binding protein-2; RanGAP, Ran GTPase-activating protein; RANK, receptor activator of NF-κB, Rub1, related to ubiquitin-1; SCF, SKP1–CUL1–F-box; SENP, Sentrin-specific protease; SKP, S-phase-kinase-associated protein; Smt3/SMT3, suppressor of *MIF2* mutations; STAT, signal transducer and activator of transcription; SUMO, small ubiquitin-like modifier; TF, transcription factor; TOR, target of rapamycin; TPPII, tripeptidyl peptidase II; TRAF, tumour necrosis factor (TNF)-receptor-associated factor; UBA domain, ubiquitin-associated domain; UBA/UBE, ubiquitin-activating enzyme; Ub-AMC, ubiquitinyl-7-amino-4-methylcoumarin; UBC, ubiquitin-conjugating enzyme; Ub-H, ubiquitin aldehyde; Ubl/UBL, ubiquitin-like protein; UCRP, ubiquitin cross-reactive protein; ULA/ULE, ubiquitin-like activating enzyme; Urm1/URM1, ubiquitin-related modifier-1; USP, ubiquitin-specific protease.

Sponsored by BIOMOL International, L.P. Designed by Fiona Montgomery and Judith Shadwell. Edited by Rachel Smallridge.