Research Article, *Biochemistry*

July 7, 2020

Supporting Information (SI) for:

The ATF3 Transcription Factor Is a Short-Lived Substrate of the Arg/N-Degron Pathway

Tri T. M. Vu and Alexander Varshavsky*

Division of Biology and Biological Engineering,
California Institute of Technology, Pasadena, CA 91125, USA

*To whom correspondence should be addressed: avarsh@caltech.edu

Content:

Supplementary Figures S1-S7 and legends to them.

Tables S1 and S2.

SI References.

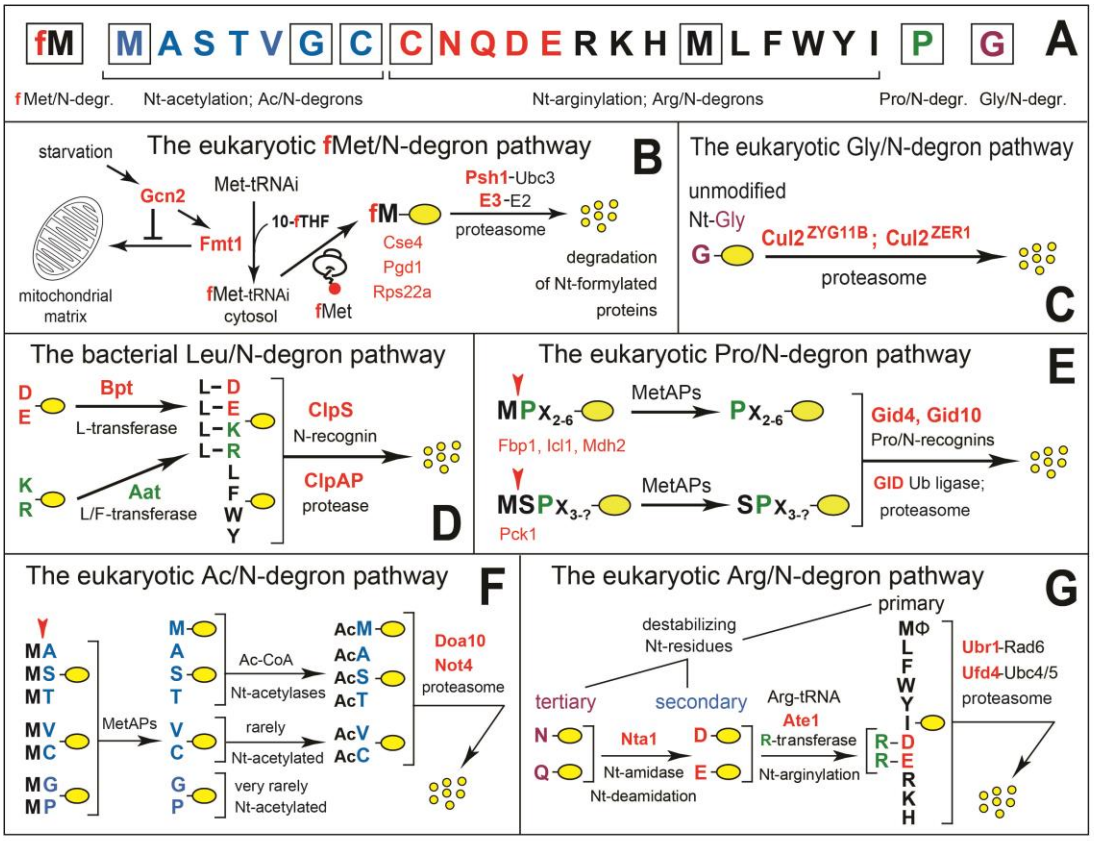

Figure S1. N-degron pathways. N-terminal (Nt) residues are indicated by single-letter abbreviations. A yellow oval denotes the rest of a protein substrate. (A) Twenty amino acids of the genetic code are arranged to delineate specific N-degrons. Nt-Met is cited thrice, since it can be recognized by the Ac/N-degron pathway (as Nt-acetylated Ac-Met), by the Arg/N-degron pathway (as unacetylated Nt-Met), and by the fMet/N-degron pathway (as Nt-formylated fMet). Nt-Cys is cited twice, since it can be recognized by the Ac/N-degron pathway (as Nt-acetylated Cys) and by the Arg/N-degron pathway (as arginylatable Nt-Cys sulfinate or Nt-Cys-sulfonate, formed in multicellular eukaryotes but apparently not in unstressed *S. cerevisiae*). (B) The eukaryotic (*S. cerevisiae*) fMet/N-degron pathway. 10-fTHF, 10-formyltetrahydrofolate. (C) The bacterial (*E. coli*) fMet/N-degron pathway. (D) The bacterial (*Vibrio vulnificus*) Leu/N-end rule pathway. (E) The eukaryotic (*S. cerevisiae*) Pro/N-degron pathway. (*F*) The eukaryotic (*S. cerevisiae*) Ac/N-degron pathway. (G*)* The eukaryotic (*S. cerevisiae*) Arg/N-degron pathway. See Introduction for references and other details.

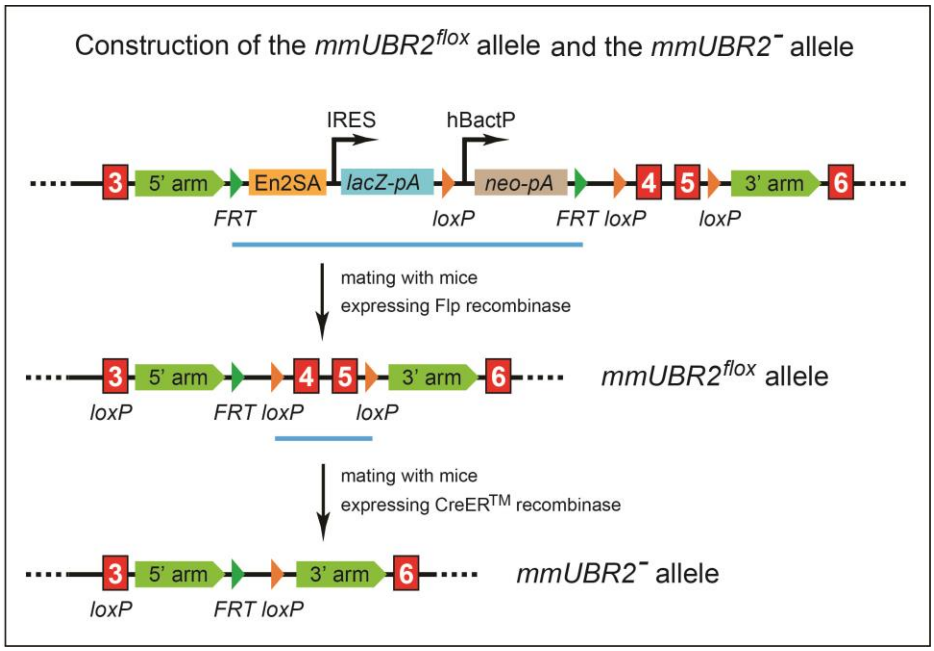

Figure S2. Construction of the mouse *mmUBR2*^{flox} allele and the *mmUBR2*$^{-}$ allele. Shown is a map of the ~30-kb 5'-proximal region of the ~90 kb *mmUBR2* gene that encompasses a gene segment from exon 3 to exon 6 (vertical red rectangles). Sizes of depicted genetic elements are not to scale. Top diagram: the *mmUBR2*$^{tm1a(KOMP)Mbp}$ allele, from mouse embryonic stem (ES) cells that were produced by the NIH-supported Knockout Mouse Project (KOMP) and was obtained from the KOMP Repository (www.komp.org). We used these ES cells to generate a mouse strain *mmUBR2*$^{tm1a(KOMP)Mbp/+}$ bearing this allele (see Results). Depicted from left to right are: the 5'-arm region of sequelogy (near-identity) to the KOMP's targeting vector; *FRT* recombination site, recognized by the bacterial Frt recombinase; mouse *En2* intron and mRNA splice acceptor (*En2SA*)[1]; an internal ribosomal entry site (IRES) from encephalomyocarditis virus; *E. coli lacZ* gene, encoding β-galactosidase and containing SV40 polyadenylation signal (*lacZ-pA*); *loxP* recombination site, recognized by the bacterial Cre recombinase; human β-actin promoter (*hBactP*); neomycin resistance cassette with SV40 polyadenylation signal (*neo-pA*); the second *FRT* recombination site; the second *loxP* recombination site; and the 3'-arm region of sequelogy (near-identity) to the KOMP's targeting vector that encompasses *mm*UBR2 exons 4 and 5, with the third *loxP* recombination site downstream of exon 5. Middle diagram: the *mmUBR2*^{flox} allele was generated by mating *mmUBR2*$^{tm1a(KOMP)Mbp/+}$ mice with *Gt(ROSA)26Sor*$^{tm1(FLP1)Dym}$ mice that expressed Flp recombinase (see Results). These matings led to the Flp-mediated excision of the *lacZ-pA-neo-pA* DNA segment between exons 3 and 4 of the *mmUBR2* gene. Bottom diagram: the *mmUBR2*$^{-}$ allele was generated by mating homozygous *mmUBR2*$^{flox/flox}$ mice and *CaggCreER*^{TM} mice that expressed CreER^{TM} recombinase. Subsequent intraperitoneal injections of tamoxifen (TM) led to the Cre-mediated excision of the "floxed" DNA segment (indicated by a blue line) that encompassed exons 4 and 5 of the *mmUBR2* gene.

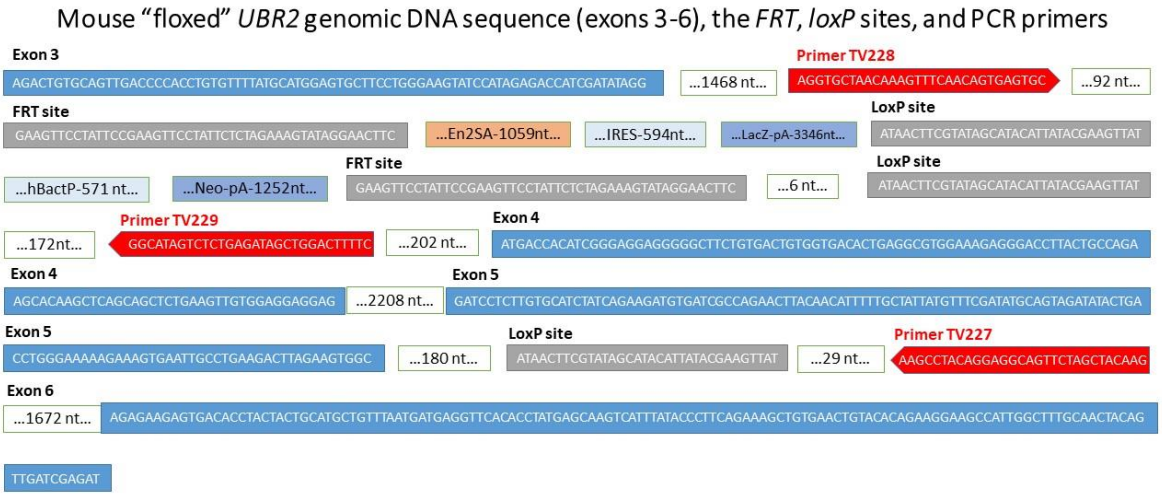

Figure S3. Mouse genomic nucleotide sequences encompassing the *mmUBR2*$^{tm1a(KOMP)Mbp}$ allele described in Figure S2. Shown are some nucleotide sequences of the *mmUBR2* exons 3 to 6 (in blue), the *FRT* and *loxP* sites, and locations of PCR primers. The names of *FRT* and *loxP* sites are in black, and their nucleotide sequences are in grey; oligonucleotide PCR primers (see Results and Table S2) are in red.

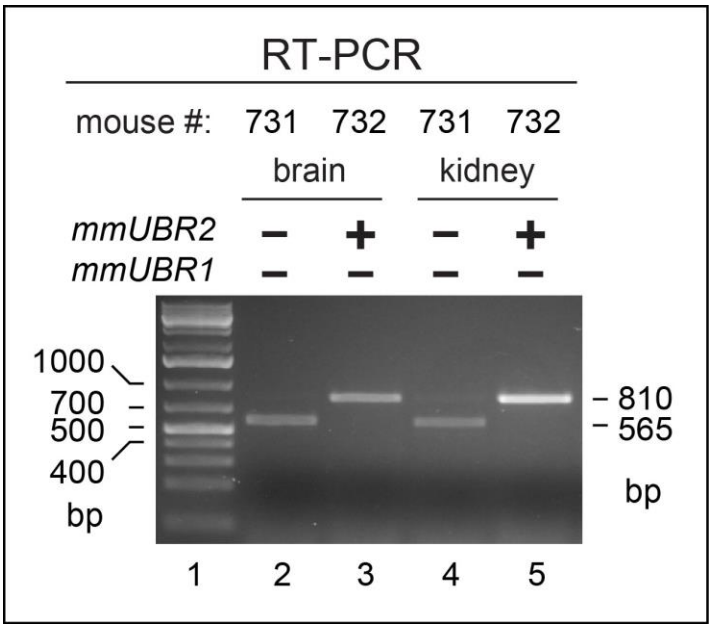

Figure S4. RT-PCR analyses with brain and kidney RNA/cDNA preparations as well as primers TV507 and TV510 (Table S2), which produced a PCR-amplified 565 bp DNA fragment for the excision-inactivated *mmUBR2*$^{-}$ allele and a 810 bp fragment for the initial ("floxed") *mmUBR2*^{flox} allele (see Figures S2 and S3, and Results).

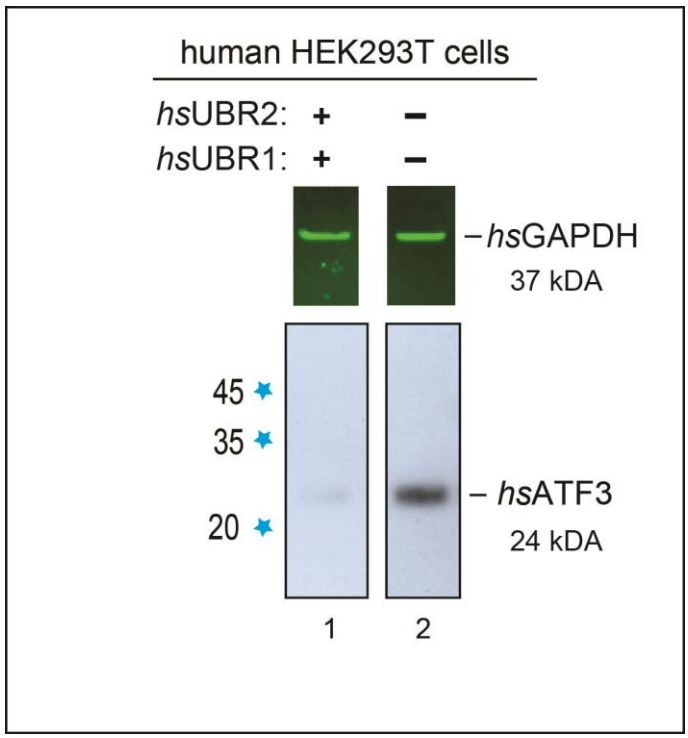

Figure S5. Immunoblotting(IB)-based comparisons of the levels of endogenous, untagged *hs*ATF3 in wild-type versus [*hsUBR1*$^{-/-}$ *hsUBR2*$^{-/-}$] HEK293T human cell lines. Lane 1, *hs*ATF3 in wild-type HEK293T cells. Lane 2, same as in lane 1 but in [*hsUBR1*$^{-/-}$ *hsUBR2*$^{-/-}$] HEK293T cells. Upper panels show the corresponding (determined by IB) levels of GAPDH (a loading control). The sizes of molecular mass markers are indicated on the left. See also Materials and Methods.

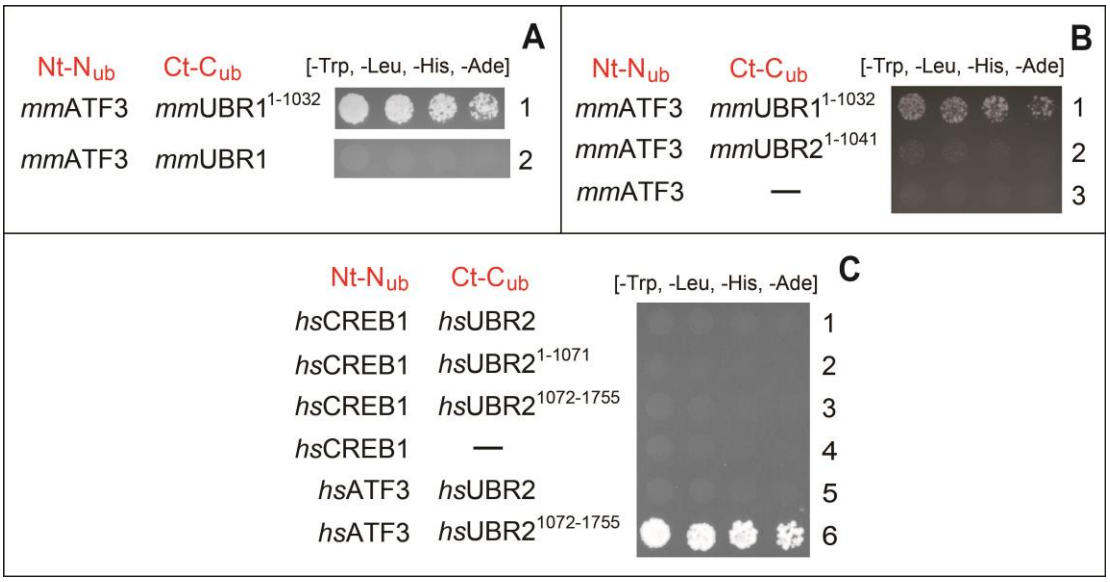

Figure S6. Split-ubiquitin protein interaction assays with transcription factors ATF3 and CREB1 versus UBR1 and UBR2 E3s. See Figure 5A for the design of split-Ub assays. (A) Row 1, mouse *mm*ATF3 vs. Nt-fragment of *mm*UBR1 (*hs*UBR1$^{1-1032}$). Row 2, *mm*ATF3 vs. full-length *mm*UBR1. (B) Row 1, *mm*ATF3 vs. Nt-fragment of *mm*UBR1 (*hs*UBR1$^{1-1032}$). Row 2, *mm*ATF3 vs. Nt-fragment of *mm*UBR2 (*hs*UBR1$^{1-1041}$). Row 3, *mm*ATF3 vs. vector alone. (C) Row 1, *hs*CREB1 vs. full-length *hs*UBR2. Row 2, *hs*CREB1 vs. Nt-fragment of *hs*UBR2 (*hs*UBR1$^{1-1071}$). Row 3, *hs*CREB1 vs. Ct-fragment of *hs*UBR2 (*hs*UBR1$^{1072-1755}$). Row 4, *hs*CREB1 vs. vector alone. Row 5, *hs*ATF3 vs. full-length *hs*UBR2. Row 6, *hs*ATF3 vs. Ct-fragment of *hs*UBR2 (*hs*UBR2$^{1072-1755}$).

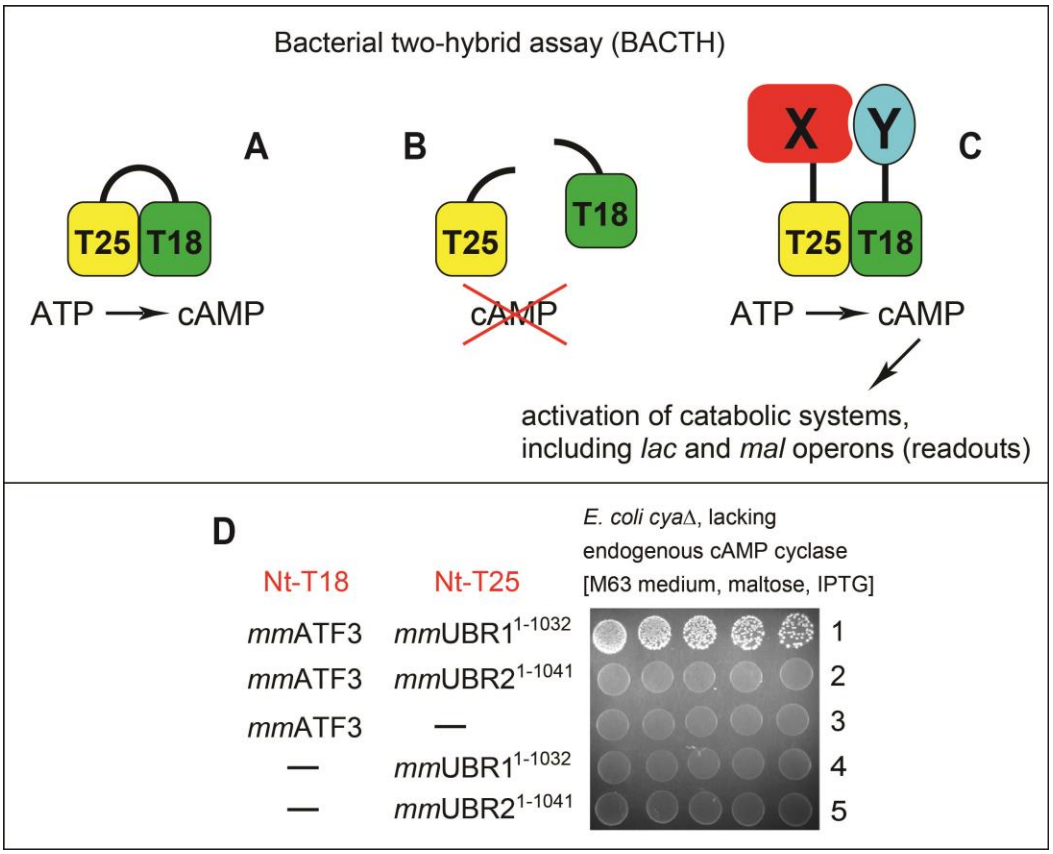

Figure S7. Use of *E. coli*-based BACTH (bacterial two-hybrid) protein interaction assay to analyze some of the binding patterns that were observed through *S. cerevisiae*-based split-Ub assays. (A) Design of the BACTH assay[2] (see Results, and Materials and Methods). (B) Row 1, mouse *mm*ATF3 vs. Nt-fragment of *mm*UBR1 (*hs*UBR1$^{1-1032}$). Row 2, *mm*ATF3 vs. Nt-fragment of *mm*UBR2 (*hs*UBR1$^{1-1041}$). Row 3, *mm*ATF3 vs. vector alone. Row 4, Nt-fragment of *mm*UBR1 (*hs*UBR1$^{1-1032}$). vs. vector alone. Row 5, Nt-fragment of *mm*UBR2 (*hs*UBR1$^{1-1041}$). vs. vector alone.

Table S1. Plasmids used in this study.

Name	Description	Source
pTV462	pX459 based, target *hsUBR1* exon 5	This paper
pTV463	pX459 based, target *hsUBR2* exon 5	This paper
pBW365	flag-mDHFR-ha-Ub-**R**-mRGS4-flag	[3]
pBW366	flag-mDHFR-ha-Ub-**V**-mRGS4-flag	[3]
pDHB1	Split-Ub vector, Ost4-C_{Ub}-LexA-VP16	MoBiTec P01001DS
pPR3-N	Split-Ub vector, N_{UbG} vector	MoBiTec P01001DS
pAI-Alg5	Split-Ub control, N_{UbI}-Alg5	MoBiTec P01001DS
pDL2-Alg5	Split-Ub control, N_{UbG}-Alg5	MoBiTec P01001DS
pTV356	Ost4-C_{Ub}-*mm*UBR1$^{1-1032}$-LexA-VP16	This paper
pTV359	Ost4-C_{Ub}-*mm*UBR2$^{1-1041}$-LexA-VP16	This paper
pTV368	N_{UbG}-*mm*ATF3	This paper
pTV471	Ost4-C_{Ub}-*hs*UBR1-LexA-VP16	This paper
pTV472	Ost4-C_{Ub}-*hs*UBR1$^{1-1059}$-LexA-VP16	This paper
pTV473	Ost4-C_{Ub}-*hs*UBR1$^{1060-1749}$-LexA-VP16	This paper
pTV679	Ost4-C_{Ub}-*hs*UBR2-LexA-VP16	This paper
pTV680	Ost4-C_{Ub}-*hs*UBR2$^{1-1071}$-LexA-VP16	This paper
pTV681	Ost4-C_{Ub}-*hs*UBR2$^{1072-1755}$-LexA-VP16	This paper
pTV466	N_{UbG}-*hs*ATF3	This paper
pTV643	N_{UbG}-*hs*CREB1	This paper
pTV645	N_{UbG}-*hs*REST	This paper
pKT25	BATCH vector, T25-MCS	EUROMEDEX EUK001
pKT25-zip	BATCH control, T25-GCN4-leucine-zipper	EUROMEDEX EUK001
pUT18C	BATCH vector, T18-MCS	EUROMEDEX EUK001
pUT18C-zip	BATCH control, T18-GCN4-leucine-zipper	EUROMEDEX EUK001
pTV672	T25-ha-*mm*UBR1$^{1-1032}$	This paper
pTV673	T25-ha-*mm*UBR2$^{1-1041}$	This paper
pTV674	T18-ha-*mm*ATF3	This paper

Table S2. Oligonucleotide primers used in experiments with [*CaggCreER*^{TM} *mmUBR1*$^{-/-}$ *mmUBR2*$^{flox/flox}$] and other mouse strains.

Name	Nucleotide Sequence (5'-3')	Use
TV230	GAGATAGGAAACTGCATGCGCTGC	*mmUBR1*$^{-}$ genotyping
TV231	GCCACTTGTGTAGCGCCAAGTGCCAG	*mmUBR1*$^{-}$ genotyping
TV232	CAAGAGTGCAACAGTTACCACATG	*mmUBR1*$^{-}$ genotyping
TV233	CTACTGCATGCTGTTTAATGATGAG	*mmUBR2*$^{-}$ genotyping
TV234	CCAGCTCATTCCTCCCACTCATGATC	*mmUBR2*$^{-}$ genotyping
TV235	GGAGGTAGAAACATGCAAATCTCTG	*mmUBR2*$^{-}$ genotyping
TV227	CTTGTAGCTAGAACTGCCTCCTGTAGGCTT	*mmUBR2*^{flox} genotyping
TV228	AGGTGCTAACAAAGTTTCAACAGTGAGTGC	*mmUBR2*^{flox} genotyping
TV229	GAAAAGTCCAGCTATCTCAGAGACTATGCC	*mmUBR2*^{flox} genotyping
TV507	GAGATCGCAGGGAGATGGCTGCAAG	*mm*UBR2 RT-PCR
TV510	GCTGGTGTTCCTCACAATGACTGTC	*mm*UBR2 RT-PCR
TV969	AGAAGGAGAAGACGGAGTGC	*hs*ATF3 RT-qPCR
TV970	TCTGAGCCTTCAGTTCAGCA	*hs*ATF3 RT-qPCR
TV766	CACCG ACGCTGTCCGTTGAATGAAG	pTV462 construction
TV767	AAAC CTTCATTCAACGGACAGCGT C	pTV462 construction
TV768	CACCG ATTGCCAGCAGATTTAGAGA	pTV463 construction
TV769	AAAC TCTCTAAATCTGCTGGCAAT C	pTV463 construction

SI References

(1) Skarnes, W. C., Auerbach, B. A., and Joyner, A. L. (1992) A gene trap approach in mouse embryonic stem cells: the lacZ reported is activated by splicing, reflects endogenous gene expression, and is mutagenic in mice *Genes Dev. 6*, 903-918.

(2) Karimova, G., Pidoux, J., Ullman, A., and Ladant, D. (1998) A bacterial two-hybrid system based on reconstituted signal transduction pathway. *Proc. Natl. Acad Sci. USA 95*, 5752-5756.

(3) Wadas, B., Piatkov, K. I., Brower, C. S., and Varshavsky, A. (2016) Analyzing N-terminal arginylation through the use of peptide arrays and degradation assays. *J. Biol. Chem. 291*, 20976-20992.