Supplementary Information

Protein Purification. For the SUMO-Josephin fusions, lysis, clarification and subtractive purification were performed as described previously, expect for slight differences in the buffers used (Weeks et al, 2007). For both ataxin-3 and ATXN3L, a single glycine residue was inserted between the Ud1 cleavage site and the Josephin domain's N-terminus (immediately prior to the start methionine). This glycine residue was included to increase the efficiency of cleavage by the SUMO-specific protease Ud1.

For ataxin-3 and ATXN3L, cells were lysed and loaded onto a HiTrap Chelating HP column in 50 mM sodium phosphate, 250 mM NaCl, 12.5 mM imidazole, 5 mM BME, pH 8, and eluted with 250 mM imidazole in the same buffer. Josephins-1 and -2 were isolated in the same buffers, except containing 25% w/v glycerol. Following cleavage and subtractive purification, the individual Josephin domains were further purified by ion exchange chromatography followed by gel filtration. The latter step was performed in 20 mM Tris-Cl, 250 mM NaCl, 5 mM DTT, pH 7.0. The desired fractions were pooled and concentrated using a YM10 concentrator device (Millipore; Danvers, MA, USA) and the proteins aliquoted, flash-frozen, and stored at -80° C.

For the various monomeric ubiquitin constructs, cell pellets were resuspended in phosphate-buffered saline, pH 7.4, containing 10 mM magnesium chloride and 1 μg/ml DNAse and RNAse. Cells were lysed with an Emulsiflex-C5 homogenizer (Avestin, Ontario, Canada) at room temperature, using conditions previously described (Weeks et al, 2007). Following clarification at 18000g, 0.5% (v/v) perchloric acid was added to the cell lysate. Following 30 minutes stirring at room temperature, the precipitated slurry was clarified again by a centrifugation step at 18000g. For all the ubiquitin constructs, except the hexahistidine tagged construct, the supernatants were dialysed against 20mM sodium acetate pH 5.2 and loaded onto HiTrap SP HP column equilibrated in the same buffer. The protein was eluted with a 0-30% sodium chloride gradient over 20 column volumes at the same pH. Following concentration of pooled fractions using an YM3 Centricon device, aliquots of the individual proteins were flash frozen and stored at -80° C until needed. For the Ub-His_6 construct, the protein was first dialysed against 50 mM sodium phosphate, 250 mM NaCl, 20 mM imidazole, pH 8, and then loaded onto a HiTrap Chelating HP column. Following elution with an imidazole gradient the protein was

further purified by cation exchange chromatography and concentrated as above. For cleavage assays, or for the synthesis of ubiquitin chains, the appropriate aliquots were rapidly thawed and the samples dialysed against water before determination of concentration by absorbance.

Synthesis of the di- and poly-ubiquitin chains was carried out using published procedures (Pickart & Raasi, 2005). The recombinant enzymes required for synthesis of the chains were all isolated to homogeneity using standard methods for IMAC and ion exchange. Following synthesis of the K48- and K63-linked poly-ubiquitin chains, the product was used directly in cleavage assays without any further purification. K48- and K63-linked di-ubiquitin chains were prepared and purified as described (Weeks et al, 2009).

The production of the ubiquitin-thioester from an intein fusion and the subsequent formation of ubiquitin$_{1-75}$-chloroethylamine were carried out based on published methods (Ovaa et al, 2005; Wilkinson et al, 2005). Cells expressing the ubiquitin$_{1-75}$-*M. xenopii* intein fusion were lysed in IMAC loading buffer (50 mM sodium phosphate, 250 mM NaCl, 12.5 mM imidazole) containing 10 mM magnesium chloride and 1 μg/ml DNAse and RNAse. Cells were passed through an iced Emulsiflex-C5 homogenizer as previously described (Weeks et al, 2007). Following clarification by centrifugation the protein was captured on a HiTrap Chelating HP column and eluted with 20 mM sodium acetate, 500 mM NaCl. Eluted fractions were immediately neutralised by the addition of 1M HEPES pH 8 to a final concentration of 30 mM. The concentration of the intein fusion protein was determined using absorbance and the protein diluted to 0.2 mM with 20 mM MES pH 6.5. EDTA was added to final concentration of 1 mM. To this solution powdered MESNA (Sigma-Aldrich) was added to a final concentration of 500 mM. The thioester substitution reaction was allowed to proceed overnight at room temperature. Following the reaction the protein was dialysed against two changes of 20 mM MES pH 6.5, 250 mM NaCl at 4° C for a total of 3 hours. The dialysed protein was then reloaded onto the same IMAC column to remove the intein and any uncleaved fusion protein. The flowthrough was concentrated using an stirred cell device (Millipore) with a YM3 membrane. The resulting ubiquitin thioester was either used immediately or stored at -80° C for a maximum of 1 week. To prepare ubiquitin$_{1-75}$-chloroethylamine, for every 500 μl of protein at a concentration of 4 - 5 mg/ml, 0.8 mmol of powdered 2-chloroethylamine was added. To this reaction 100 μl of 2 M NaOH were added and solution mixed. Following 10 to 15 minutes incubation at room temperature, 200 μL of 1M HCl were added to neutralise the reaction. The solution was then

dialysed using a 3500 MWCO Slide-a-Lyzer (Pierce) against 20 mM sodium acetate pH 5.2. The dialysed sample was then loaded onto a 5 ml Hitrap SP column, equilibrated in the same buffer, and eluted with a 0-30% NaCl gradient, which resulted in baseline separation of the halogenated ubiquitin from the hydrolysed thioester. The ubiquitin$_{1-75}$-chloroethylamine was concentrated and then used immediately for conjugation with purified ATXN3L protein.

For the formation of the complex, the ATXN3L was isolated by subtractive purification as above. Following the second IMAC step the protein was concentrated to 10 mg/ml and then dialysed against 20 mM Tris pH 7.5, 1 mM EDTA, 10% w/v glycerol, 5 mM DTT. The purified ubiquitin$_{1-75}$-chloroethylamine was added to the concentrated ATXN3L in slight molar excess, in a mixture containing an additional 50 mM Tris pH 7.5. The conjugation reaction was incubated for approximately 72 hours at RT until the formation of the complex had reached almost 95% based on SDS-PAGE analysis. The reaction was then loaded directly on a 5 ml HiTrap Q HP column equilibrated in 20 mM Tris pH 7.0, 5 mM DTT. Protein was eluted using a 0 to 40% NaCl gradient over 20 column volumes. Fractions corresponding to the conjugated protein were concentrated and then loaded onto an S200 Sephacryl 16/60 column equilibrated in 20 mM Tris pH 7.0, 250 mM NaCl, 0.5 mM EDTA, 5 mM DTT. Fractions were pooled and concentrated to approximately 40 mg/ml using a YM10 concentrator. Aliquots of the purified protein were flash-frozen and stored at -80° C until needed. The preparation of the selenomethionine-labelled ATXN3L-ubiquitin conjugate was carried out in exactly the same manner.

Table S-I. Surface areas buried during Josephin-ubiquitin complex formation for the four copies of the complex in the crystal asymmetric unit.

Buried surface area (Å^2)[a]	**A:B**	**C:D**	**E:F**	**G:H**
ATXN3L	1236	1202	1213	1098
Ubiquitin	1455	1409	1453	1322

[a] Buried surface areas differ for the ATXN3L and ubiquitin components of each complex because the C-terminus of ubiquitin is buried in the ATXN3L active site, which is largely inaccessible to solvent even in the unliganded enzyme.

Table S-II. Polar interactions between ubiquitin and the helical hairpin formed by Josephin helices 2 & 3.

ATXN3L residue	**Ubiquitin residue**	**Type of interaction**	**Distance (Å)**
Glu-57	Lys-48	Salt bridge	2.6
Gln-64	Gln-49	H bond	3.2
Glu-44	Arg-42	Salt bridge	3.1
Glu-67	Arg-72	Salt bridge	2.8/3.3
Asp-70	Arg-74	Salt bridge	2.5
Gln-78	Leu-71 (backbone amide)	H bond	3.0

Table S-III. Conservation of the ataxin-3 Josephin domain among vertebrates.

		Human	Macaque		Mouse	Platypus	Chicken	Zebrafish
		ataxin-3	ataxin-3[b]	ATXN3L	ataxin-3	ataxin-3	ataxin-3	ataxin-3[b]
Human	ataxin-3	100[a]/100	100/100	86.7/96.7	99.4/100	97.2/98.9	95.6/97.8	83.9/92.8
	ATXN3L	85.0/92.2	81.7/88.3	95.6/98.3	84.4/92.2	83.9/92.2	83.9/92.2	73.3/86.7

[a] Residues 1-180 of the Josephin domain were subjected to global pairwise alignment using the Needleman-Wunsch global alignment algorithm in EMBOSS (Rice et al, 2000); values in the table reflect percent identity/percent similarity, as determined by the program.

[b] Currently, the predicted sequences for these proteins are N-terminally truncated by a number of residues; therefore the values given may be slightly different for the full Josephin domain.

Table S-IV: Expression constructs generated and used in this work. Plasmids pETHSUL, pETHSU, pETCH and pETXSH are in house vectors, the first three have been described elsewhere (Weeks et al, 2007). The ATXN3L and the human Josephins were amplified from constructs kindly provided by Randy Pittman (Univ. of Pennsylvania). Templates for enzymes for poly-ubiquitin chain synthesis were provided by Wai-Kwan Tang (NIH/NCI) and Roger Greenberg (Univ. of Pennsylvania), except the N-terminally truncated *H. sapiens* E1b, which was amplified from a cDNA clone (MGC:4781 IMAGE:3543021), and YUH1, which was amplified from *S. cerevisiae* genomic DNA.

Target	**Primers (5'-3')**	**Plasmid Backbone**	**Cloning Method**
Ataxin-3	1: AGATTGGTGGCGGTATGGAGTCCATCTTCCACGAG 2: GAGGAGAGTTTAGACTATTATTTTGGTCGATGCATCTGTTGG	pETHSUL	LIC
ATXN3L	1: AGATTGGTGGCGGTATGGATTTCATCTTTCATGAG 2: GAGGAGAGTTTAGACTATTATTTTGGTGTATCCATCTCTTCGA	pETHSUL	LIC
Josephin-1	1: TTTTTTGGTCTCTCATGAGTTGTGTGCCATGGAAAGG 2: TTTTTTAAGCTTCGTCTCTATTACACATCGGTCCTCCAACTCTG	pETUbHSU[a]	BsaI/HindIII digestion and ligation.
Josephin-2	1: TTTTTTGGTCTCTCATGTCCCAGGCCCCGGGAGCACAG 2: TTTTTTAAGCTTCGTCTCTATTAGTCTGTCCGCAGCCAGCTGC	pETUbHSU[a]	BsaI/HindIII digestion and ligation.
Ubiquitin	1: TTTTTTCATATGCAGATCTTCGTGAAGACCC 2: TTTTTTGAATTCTATTACCCACCTCTGAGACGGAGCA	pRSETA	NdeI/EcoRI Ligation
Ubiquitin-intein	1: GTGGTGGTGGTGGTGCCCGGGGCGGATGGTGTATCGGACATA 2: AGTGCATCTCCCGTGATACAACCTCTGAGACGGAGCACCA	pETXSH	SLIC
Ubiquitin D77	Kindly provided by S. Raasi	pET3	N/A
Ubiquitin K48C	Kindly provided by S. Raasi	pET3	N/A
Ubiquitin K63R	Phospho-TCAGCGCGAGTCTACCCTGCACCTGGT Phospho-ATATTGTAGTCAGACAGGGTGCGT	pRSETA	Mutagenesis
Ubiquitin-His_6	1: TAATACGACTCACTATAGGG 2: TTTTTTCCCGGGACCGCCCCCACCACGGAGACGCAGCACCA	pETCH	XbaI/XmaI digestion and ligation.
E1b	1: TTAAGAAGGAGATATACCATGGCCAAGAACGGCAGTGAA 2: GTGGTGGTGGTGGTGCCCGGGGCGGATGGTGTATCGGACATA	pETCH	SLIC
E225K	1: TTAAGAAGGAGATATACCATGGCCAACATCGCGGTGCA 2: GTGGTGGTGGTGGTGCCCGGGGTTACTCAGAAGCAATTCTGTTG	pETCH	SLIC
Mms2	1: AGATTGGTGGCATGTCAAAAGTACCAAGAAATTTTAG 2: GAGGAGAGTTTAGACTATTAAAAGGTTTCTCCTTCCTTCGGTT	pETHSUL	LIC

Table S-IV, *continued*

Ubc13	1: TTAAGAAGGAGATATACCATGGCATCATTACCCAAGAGAAT 2: GTGGTGGTGGTGGTGCCCGGGCTCGGGTTTCTTCTTTGCATAC	pETCH	SLIC
YUH1	1: TTTTTTCCATGGGTCACCACCATCACCACCATAGCGGAGAAAA TCGTGCTGTG 2: TTTTTTAAGCTTATTCCCAATTAGGGCCCAATC	pETCH	NcoI/HindIII Ligation

[a] The ubiquitin coding region was inserted into the NcoI site, upstream of the SUMO fusion sequence.

Table S-V. Primers used for site-directed mutagenesis of Ataxin-3. pETHSUL-Ataxin-3 was used as a template.

Ataxin-3 mutant construct	**Primers (5'-3')**
S12F	1: GAAGGCTTCCTTTGTGCGCAACATTGC 2: GCAATGTTGCGCACAAAGGAAGCCTTC
G67E	1: ACAGCAGCCTTCTGAAAATATGGATGACAG 2: CTGTCATCCATATTTTCAGAAGGCTGCTGT
L93H	1: TTTAGAACTAATCCACTTCAACAGTCCAGA 2: TTTAGAACTAATCCACTTCAACAGTCCAGA
E118Q	1: ATATGCAATTATAAGCAGCACTGGTTTACAGTT 2: AACTGTAAACCAGTGCTGCTTATAATTGCATAT
R59L	1: GTGAAGATTATCGCACTGTTTTACAGCAGCCTTC 2: GGCTGCTGTAAAAACGTCAGATAATCTTCACTAG
T60A	1: CTAGTGAAGATTATCGCGCGTTTTTACAGCAGCC 2: GGCTGCTGTAAAAACGCGCGATAATCTTCACTAG
R59L/T60A	1: ACTAGTGAAGATTATCTGGCGTTTTTACAGCAGCC 2: GGCTGCTGTAAAAACGCCAGATAATCTTCACTAGT
S12F+R59L/T60A	S12F primers were used with the pETHSUL-R59L/T60A mutant

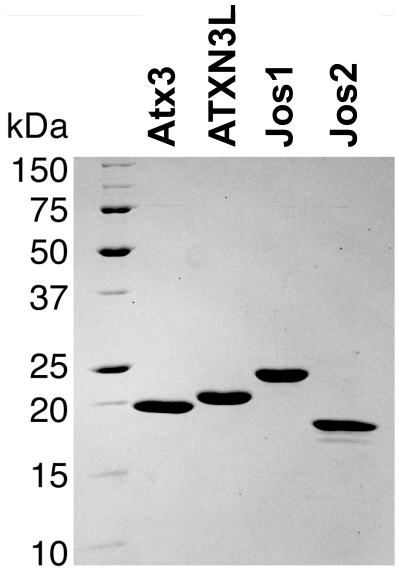

Figure S1. Purified Josephin domains from the four human Josephin-domain-containing proteins, ataxin-3 (Atx3), the ataxin-3-like protein (ATXN3L), Josephin-1 (Jos1), and Josephin-2 (Jos2). Shown is an SDS-PAGE gel, stained with Coomassie Brilliant Blue. The unlabeled lane at left contains molecular weight markers, with molecular weight values being listed at the left of the gel.

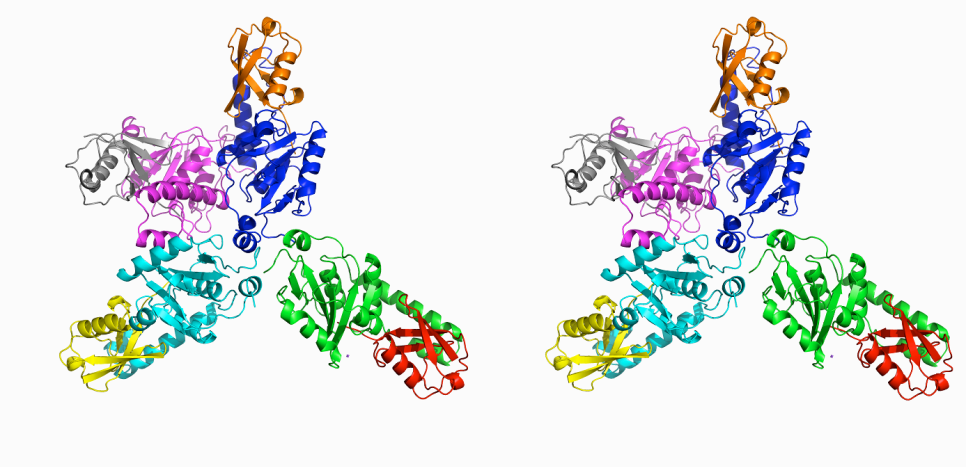

Figure S2. Stereo view of the contents of the crystal asymmetric unit. The vantage point is such that the viewer is looking along the approximate 3-fold rotation axis the relates chains A:B, C:D, and E:F. Color scheme: Chain A (ATXN3L), green; chain B (ubiquitin), red; chain C (ATXN3L), cyan; chain D (ubiquitin), yellow; chain E (ATXN3L), blue; chain F (ubiquitin), orange; chain G (ATXN3L), magenta; chain H (ubiquitin), gray.

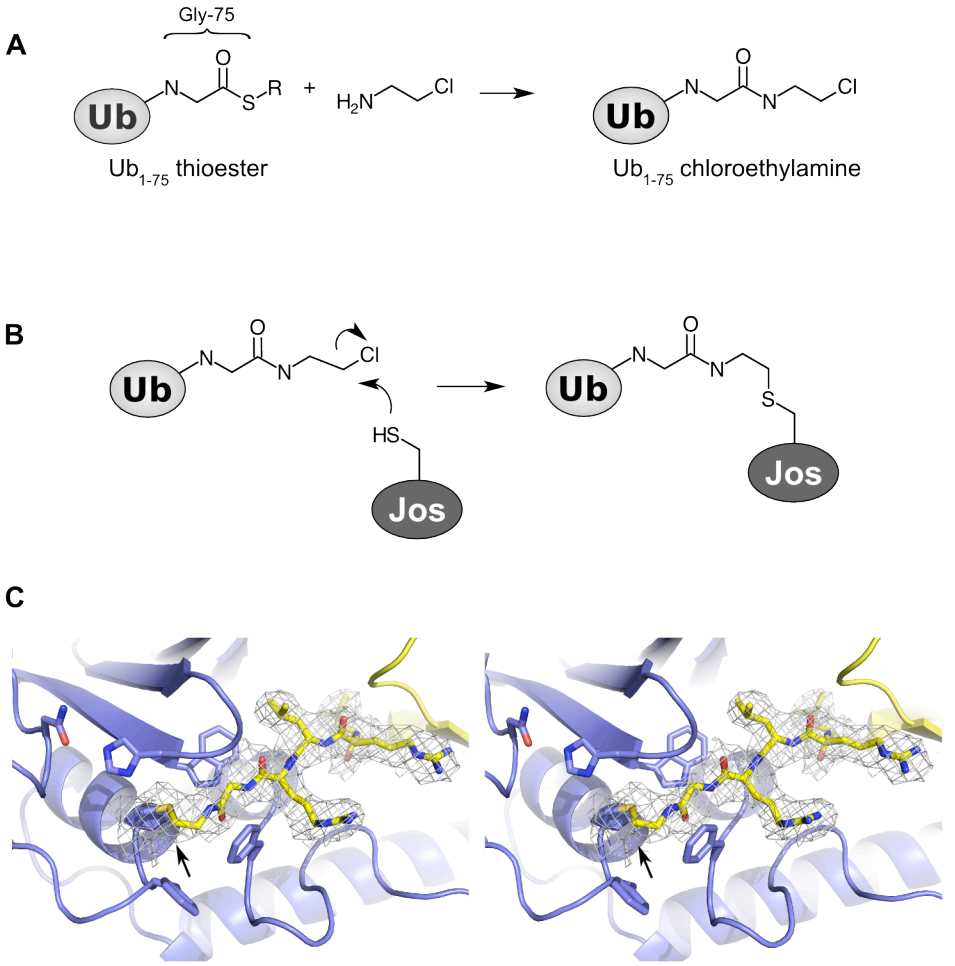

Figure S3. (**A**) Scheme detailing the production of the ubiquitin$_{1-75}$-chloroethylamine active-site-directed reagent. Residues 1-75 of ubiquitin are produced as a C-terminal thioester, using intein chemistry, after which the adduct with 2-chloroethylamine is formed under basic conditions. (**B**) The purified ubiquitin$_{1-75}$-chloroethylamine is added to the Josephin domain, upon which the active site cysteine nucleophile attacks the 2-position of the chloroethylamine group; loss of chlorine generates a stable covalent complex between the Josephin and ubiquitin. (**C**) Stereo view showing the covalent Josephin-ubiquitin linkage in the ATXN3L structure. The Josephin domain is shown in blue, and ubiquitin in yellow; the orientation of this figure is similar to that of Figure 2B. 2Fo-Fc electron density is shown for the C-terminal residues of ubiquitin and the active site cysteine (Cys-14); the position of the covalent link is marked by the arrow. Also shown are the other two members of the catalytic triad, Asn-134 and His-119. Additionally, three aromatic residues that form a hydrophobic enclosure around the active center are shown (Phe-12, Phe-74, and Trp-120).

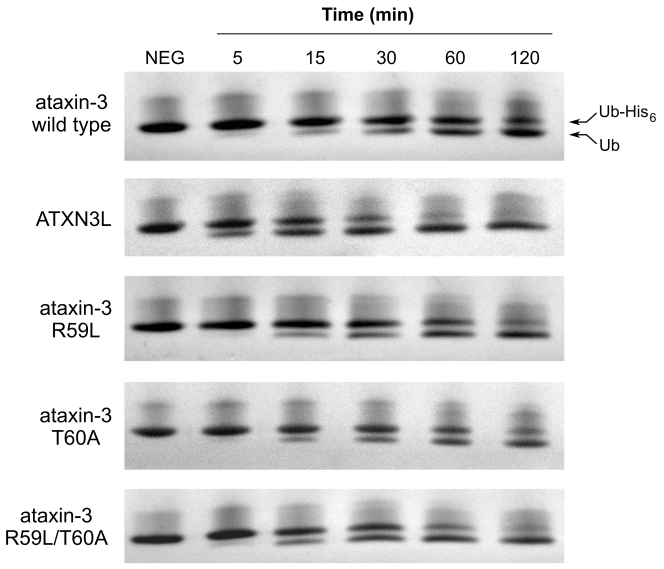

Figure S4. Effects upon catalytic efficiency of the ataxin-3 mutations at amino acids 59 and 60. Shown are time courses for cleavage reactions carried out using the various Josephin domains specified, using the Ub-His_6 substrate. NEG, negative control (no enzyme added). Reactions were carried out at room temperature.

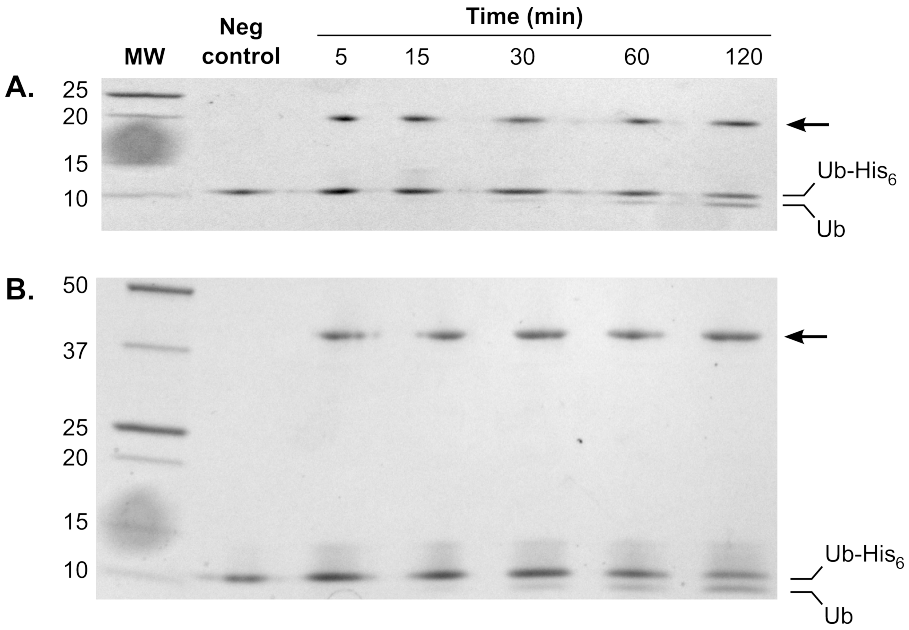

Figure S5. Time course of cleavage of the Ub-His_6 substrate by the ataxin-3 Josephin domain and full-length ataxin-3. Panel (A), Ataxin-3 Josephin domain; panel (B), full-length ataxin-3. The positions of the uncleaved and cleaved substrates are shown at right. The bold arrows show the position of the ataxin-3 enzymes. Sizes of the molecular weight markers are shown at left. Negative control = substrate without enzyme.

References

Ovaa H, Galardy PJ, Ploegh HL (2005) Mechanism-based proteomics tools based on ubiquitin and ubiquitin-like proteins: synthesis of active site-directed probes. *Methods Enzymol* **399:** 468-478

Pickart CM, Raasi S (2005) Controlled synthesis of polyubiquitin chains. *Methods Enzymol* **399:** 21-36

Rice P, Longden I, Bleasby A (2000) EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet* **16**(6)**:** 276-277

Weeks SD, Drinker M, Loll PJ (2007) Ligation independent cloning vectors for expression of SUMO fusions. *Protein Expr Purif* **53**(1)**:** 40-50

Weeks SD, Grasty KC, Hernandez-Cuebas L, Loll PJ (2009) Crystal structures of Lys-63-linked tri- and di-ubiquitin reveal a highly extended chain architecture. *Proteins* **77**(4)**:** 753-759

Wilkinson KD, Gan-Erdene T, Kolli N (2005) Derivitization of the C-Terminus of Ubiquitin and Ubiquitin-like Proteins Using Intein Chemistry: Methods and Uses. *Methods Enzymol* **399:** 37-51