Supporting Information

Site-Specific *in Vivo* Protein SUMOylation via Translational Incorporation of a Proximity-Reactive Pyrrolysine Analogue

Yuk Hei Chan, Marianne M. Lee, Michael K. Chan*

School of Life Sciences & Center of Novel Biomaterials, The Chinese University of Hong Kong, Shatin, Hong Kong S.A.R.

*Corresponding author: Michael K. Chan, E-mail: michaelkchan88@cuhk.edu.hk, Tel: +852-39431487, Fax: + 852-2603 5646

Table of Contents

Experimental Methods

mCherry reporter assay

To evaluate the readthrough efficiency of 2-chloroacetyl-Nε-lysine (ClAcK) by PylRS and MmFAcKRS1, the relevant reporter plasmid (pETDuet-PylT-PylRS-KanR-K46UAG-mCherry-K55UAG or pETDuet-PylT-MmFAcKRS1-KanR-K46UAG-mCherry-K55UAG) was first transformed to BL21(DE3). Single colonies of respective constructs were then picked and cultured overnight in 5 mL of LB medium with appropriate antibiotics. After incubation at 30°C, 5 µL of each culture was transferred to 495 µL of M9 minimal medium with 100 µg/mL ampicillin, 0.05 mM IPTG and varying concentration (0-2 mM) of ClAcK in a 24-wells plate. The normalized mCherry reading [mCherry reading (Ex: 587 ± 5 nm; Em: 610 ± 5 nm) divided by the optical density at 700 nm] was measured with Tecan Infinite M1000 PRO.

Expression of EcADC(wt)-His_6 and EcADC-K240X/D535C-His_6

The genes for both EcADC(wt)-His_6 and EcADC-K240UAG/D525C-His_6 were cloned into the multiple cloning site 2 (MCS2) of the pETDuet-PylT-MmFAcKRS1 vector. To express the proteins, the relevant plasmids were transformed into BL21(DE3) cells. The transformed cells were then spread onto LB agar plates containing ampicillin and incubated overnight at 37 °C. Successful transformants were inoculated into 500 mL of LB medium supplemented with 100 µg/mL ampicillin the following day and incubated at 37°C with shaking at 220 rpm. The cells were harvested and resuspended in 100 mL (1/5th of the original volume) of fresh LB medium once the optical density at 600 nm (OD_{600}) reached 0.6. The cultures were supplemented with 100 µg/mL ampicillin and 0.2 mM IPTG. For EcADC-K240X/D535C-His_6, 2 mM ClAcK and 10mM nicotinamide were additionally provided. Protein expression occurred overnight at 30 °C with 220 rpm shaking.

Expression of site-specifically SUMOylated proteins in *E. coli*

For ClAcK-mediated TDG SUMOylation, relevant plasmids were first transformed into BL21(DE3) cells. The transformed cells were then spread onto LB agar plates containing ampicillin and incubated overnight at 37 °C. Successful transformants were inoculated into 500 mL of LB medium supplemented with 100 µg/mL ampicillin the following day and incubated at 37°C with shaking at 220 rpm. The cells were harvested and resuspended in 100 mL (1/5th of the original volume) of fresh LB medium once OD_{600} reached 0.6. The cultures were supplemented with 100 µg/mL ampicillin, 0.2 mM IPTG, 2 mM ClAcK and 10mM nicotinamide. Protein expression occurred overnight at 20 °C with 220 rpm shaking. ClAcK-mediated αSyn SUMOylation followed the same procedures as the TDG SUMOylation, with the exception that protein expression occurred overnight at 30 °C with 220 rpm shaking. For negative control with no reactivity to cysteine, 2 mM ClAcK is replaced with 2 mM Nε-acetyl-lysine (AcK).

Western blotting analysis of ClAcK-mediated TDG SUMOylation

Following protein expression, cells were harvested by centrifugation (8,000 rpm for 8 minutes) and resuspended in20 mL lysis buffer (50 mM Tris pH 8.0, 150 mM NaCl, 10 mM imidazole). Lysis was achieved by sonication on ice for 20 minutes (2 s on, 2 s off, 40% power). Following centrifugation at 13,000 rpm for 30 minutes, the clarified lysate was filtered through a 0.45 mm filter for SDS-PAGE.

To further enrich the His-tagged TDG, the filtrate was purified by loading onto a Ni-charged Profinity IMAC column (5 mL, Bio-Rad). After washing with 40 mM imidazole, both TDG and SUMOylated TDG were eluted with elution buffer (50 mM Tris pH 8.0, 150 mM NaCl, 300 mM imidazole). Both the clarified cell lysates and enriched protein samples were then subjected to SDS-PAGE with stained molecular weight marker (Bio-rad, Precision Plus Protein Dual Color Standards)

After SDS-PAGE analysis of the samples, proteins were transferred to a nitrocellulose membrane (Bio-Rad), blocked with 5% Blotting-Grade Blocker (Bio-Rad) in PBS with 0.1% Tween 20, and probed overnight at 4°C with either mouse anti-His Tag (Merck, 05-949, 1:2,000), or rabbit anti-SUMO-1 (Santa Cruz, sc-9060, 1:1,000) antibodies. The secondary antibodies used were either rabbit anti-mouse IgG (SouthernBiotech, 6170-05, 1:3,000) or goat anti-rabbit IgG (Bio-Rad, 170-6515, 1:3,000). Blots were imaged with ChemiDoc™ MP Imaging System using both colorimetric and chemiluminescence detection channels. Images from both channels for the same blot were merged for analysis. Band intensities were determined densitometrically using ImageJ software. The conjugation rate is calculated by

$$conjugation\ rate = \frac{I_{conjugate}}{I_{target} + I_{conjugate}}$$

, with $I_{conjugate}$ being the band intensity of the conjugate protein and I_{target} being the band intensity of the target protein (TDG).

Purification of SUMO(QTC)-TDG(112-360)-K330X

For the purification of SUMO(QTC)-TDG(112-360)-K330X, the enriched protein samples containing SUMO(QTC)-TDG(112-360)-K330X and unconjugated TDG(112-360)-K330X were buffer exchanged to 20 mM HEPES, pH 7.4 and loaded onto a pre-equilibrated High Q anion exchange column (5 mL, Bio-Rad). The column was then washed with washing buffer (20 mM HEPES pH 7.4, 50 mM NaCl) for 30 minutes, with SUMO(QTC)-TDG(112-360)-K330X subsequently eluted using a linear salt gradient from 50 mM to 200 mM NaCl over 30 minutes.

Western blotting analysis of ClAcK-mediated αSyn SUMOylation

Following protein expression, cells were harvested by centrifugation (8,000 rpm for 8 minutes) and resuspended in20 mL lysis buffer (50 mM Tris pH 8.0, 150 mM NaCl, 10 mM imidazole). Lysis was achieved by sonication on ice for 20 minutes (2 s on, 2 s off, 40% power). The lysate was then centrifuged at 13,000 rpm for 30 minutes. The supernatant was boiled for 10 minutes and centrifuged again at 13,000 rpm for 30 minutes. Finally, the clarified lysate was filtered through a 0.45 µm filter for SDS-PAGE.

After SDS-PAGE analysis of the samples, proteins were transferred to a nitrocellulose membrane (Bio-Rad), blocked with 5% Blotting-Grade Blocker (Bio-Rad) in PBS with 0.1% Tween 20, and probed overnight at 4°C with either rabbit anti-α-synuclein (Santa Cruz, sc-7011-R, 1:1,000), or rabbit anti-SUMO-1 (Santa Cruz, sc-9060, 1:1,000) antibodies. The secondary antibodies used was goat anti-rabbit IgG (Bio-Rad, 170-6515, 1:3,000). Blots were imaged with ChemiDoc™ MP Imaging System using both colorimetric and chemiluminescence detection channels. Images from both channels for the same blot were merged for analysis. Band intensities were determined densitometrically using ImageJ software. The conjugation rate is calculated by

$$conjugation\ rate = \frac{I_{conjugate}}{I_{target} + I_{conjugate}}$$

, with $I_{conjugate}$ being the band intensity of the conjugate protein and I_{target} being the band intensity of the target protein (αSyn).

Expression of GST-SAE1/SAE2, His_6-Ubc9, His_6-SUMO(KGG), and αSyn-V52K-K96R-K102R-His_6

To express GST-SAE1/SAE2, His_6-Ubc9, and His_6-SUMO(KGG), their respective plasmids (pGEX-E1, pET28-His_6-Ubc9, and pET28-His_6-SUMO(KGG)) were transformed to BL21(DE3) cells. These transformed cells were then spread onto LB agar plates containing the appropriate antibiotics (ampicillin for pGEX; kanamycin for pET28). After overnight incubation at 37°C, the successful transformants were inoculated to 250 mL of LB medium supplemented with the appropriate antibiotic (100 μg/mL). The cultures were then incubated at 37°C with shaking at 220 rpm. Once OD_{600} reached 0.6, protein expression was induced by adding 0.3 mM IPTG. The cultures were subsequently shaken at 220 rpm overnight at 25°C.

For expression of αSyn-V52K-K96R-K102R-His_6, the plasmid pETDuet-PylT-MmFAcKRS1-SUMO(QTC)-αSyn-V52K-K96R-K102R-His_6 was transformed into BL21(DE3) cells. The transformed cells were then spread onto LB agar plates containing ampicillin and incubated overnight at 37 °C. Successful transformants were inoculated into 800 mL of LB medium supplemented with 100 μg/mL ampicillin the following day and incubated at 37°C with shaking at 220 rpm. Once OD_{600} reached 0.6, protein expression was induced by adding 0.3 mM IPTG. The cultures were subsequently shaken at 220 rpm overnight at 30°C.

Purification of His-tagged proteins

After harvesting by centrifugation at 8,000 rpm for 8 minutes, cell pellets were resuspended in 20 mL lysis buffer (50 mM Tris pH 8.0, 150 mM NaCl, 10 mM imidazole) and lysed by sonication on ice for 20 minutes (2 s on, 2 s off, 40% power). Following centrifugation at 13,000 rpm for 30 minutes, the clarified lysate was filtered through a 0.45 mm filter. Since all ClAcK-bearing proteins, EcADC(wt)-His_6, His_6-Ubc9, His_6-SUMO(KGG), and αSyn-V52K-K96R-K102R-His_6 carry a His_6-tag on the C-terminus, the filtrate was purified by loading onto a pre-equilibrated Ni-charged Profinity IMAC column (5 mL, Bio-Rad) connected to an NGC (Bio-Rad) FPLC system. The flow rate was maintained at 2 mL/min during the purification process. The column was washed with washing buffer (50 mM Tris, pH 8.0, 150 mM NaCl, 40 mM imidazole) until weakly bound materials were removed (A_{280} flattened). His-tagged proteins were then eluted with elution buffer (50 mM Tris pH 8.0, 150 mM NaCl, 300 mM imidazole) and collected in 2 mL fractions. Pure fractions were identified by SDS-PAGE.

For His_6-Ubc9, His_6-SUMO(KGG), and αSyn-V52K-K96R-K102R-His_6, the pure fractions were further dialyzed to 50 mM Tris pH 8.0, 150 mM NaCl for downstream applications.

For protein samples used for intact protein mass spectrometry (SUMO(QTC)-TDG(112-360)-K330X-His_6 and TDG(112-360)-His_6 co-eluted with SUMO(QTC)), the fractions were dialyzed to 20 mM HEPES, pH 7.4 Subsequently, the dialyzed samples were further purified by anion exchange chromatography using an EconoFit Macro-Prep High Q column (5 mL, Bio-Rad) on an NGC (Bio-Rad) FPLC system. Protein samples were loaded onto the column and washed with washing buffer (20 mM HEPES, pH 7.4, 50 mM NaCl) for 30 minutes. Target proteins were eluted using a linear gradient of 50-200 mM NaCl over 30 minutes and collected in 2 mL fractions. Relevant fractions were identified by Western blot using either anti-His or anti-SUMO-1 antibodies and pooled together.

Purification of GST-SAE1/SAE2

After harvesting by centrifugation at 8,000 rpm for 8 minutes, cell pellets were resuspended in 20 mL lysis buffer (50 mM Tris pH 8.0, 150 mM NaCl, 1mM dithiothreitol) and lysed by sonication on ice for 20 minutes (2 s on, 2 s off, 40% power). Following centrifugation at 13,000 rpm for 30 minutes, the clarified lysate was filtered through a 0.45 mm filter and loaded onto pre-equilibrated glutathione resin in a gravity column. The column was allowed to rotate at 4 °C for 2 hours. After draining the lysate from the column, the resin was washed three times with lysis buffer. GST-SAE1/SAE2 was eluted with elution buffer (50 mM Tris pH 8.0, 150 mM NaCl, 1mM dithiothreitol, 10 mM reduced L-glutathione) and collected in 1 mL fractions. Pure fractions were identified by SDS-PAGE and further dialyzed to 50 mM Tris pH 8.0, 150 mM NaCl for downstream applications.

In vitro enzymatic SUMOylation

In vitro enzymatic SUMOylation of αSyn-V52K-K96R-K102R-His was carried out in 20 µL SUMOylation buffer (50 mM Tris, pH 7.7, 150 mM NaCl, 5 mM $MgCl_2$, 2 mM DTT, 1% glycerol, 5 mM ATP) in the presence of 210 pmol His-SUMO(KGG), 2.5 pmol GST-SAE2/SAE1, 60 pmol Ubc9, and 108 pmol αSyn-V52K-K96R-K102R-His. The reaction mixture was incubated at 37°C for 2 hours.

In vitro enzymatic SUMOylation of TDG(112-360)-His was carried out in 20 µL SUMOylation buffer (50 mM Tris, pH 7.7, 150 mM NaCl, 5 mM $MgCl_2$, 2 mM DTT, 1% glycerol, 5 mM ATP) in the presence of 210 pmol His-SUMO(KGG), 2.5 pmol GST-SAE2/SAE1, 60 pmol Ubc9, and 108 pmol TDG(112-360)-His. The reaction mixture was incubated at 30°C for 18 hours.

SUMO protease cleavage

Protein samples are adjusted to 5 µM in reaction buffer (50 mM Tris, pH 7.7, 150 mM NaCl, 5 mM $MgCl_2$, 2 mM DTT, 1% glycerol) and incubated with 0.4 mg/mL Ulp1 overnight at 30°C.

Electrophoretic mobility shift assay

DNA probe was designed as previously described.[1] The heteroduplex DNA (BGI) consisted of a 28mer target strand (`5'- GTG TCA CCA CTG CTC AUG TAC AGA GCT G -3'`) and a FAM-labeled complementary strand (`5'- CAG CTC TGT ACG TGA GCA GTG GTG ACA C -3'-FAM`). Binding reactions were initiated by mixing TDG(112-360)-His or SUMO(QTC)-TDG(112-360)-K330X at varying concentrations with 10 nM DNA in binding buffer (20 mM HEPES, pH 7.4, 100 mM NaCl, 5% glycerol, 1 mM EDTA, 1 mM DTT). Samples were incubated at room temperature for 30 minutes and then loaded onto a 6% native polyacrylamide gel. Electrophoresis was performed at 80V for 75 minutes. Gels were imaged using a ChemiDoc™ MP Imaging System.

Mass Spectrometry

Relevant protein bands from the SDS-PAGE gel were excised, destained with 50% MeOH/10 mM NH_4HCO_3, and dehydrated with acetonitrile. The gel pieces were subsequently vacuum-dried and incubated at 30°C with trypsin for overnight digestion. Digested peptides were extracted from gel through sonication and desalted by $ZipTip_{C18}$ before LC-MS/MS submission. LC-MS/MS was conducted on an Orbitrap Fusion Lumos Tribrid Mass Spectrometer (Thermo Fisher Scientific), and data analysis was performed using Proteome Discoverer (Thermo Fisher Scientific).

For determination of intact protein mass using ESI-MS, protein samples (SUMO(QTC)-TDG(112-360)-K330X-His_6 and TDG(112-360)-His_6 co-eluted with SUMO(QTC)) were buffer exchanged to water before sample submission. LC-MS was conducted on either an Orbitrap Fusion Lumos Tribrid Mass Spectrometer (Thermo Fisher Scientific) or an Agilent 6540 Quadrupole-Time-of-Flight Mass Spectrometer. Deconvolution of the intact protein mass spectrum was performed using either Protein Deconvolution (Thermo Fisher Scientific) or MassHunter (Agilent).

Analysis of reaction kinetics between ClAcK and glutathione by RP-HPLC

0.1M glutathione stock was freshly prepared by dissolving reduced L-glutathione in PBS. The pH of the stock solution was then adjusted to 7.4 by addition of NaOH. 1 M ClAcK stock was prepared by dissolving ClAcK in water. OPA solution was prepared by dissolving 25 mg o-phthalaldehyde (OPA) in 500 µL methanol followed by addition of 50 µL 2-mercaptoethanol.

For the reaction, 10 mM of glutathione was mixed with 2 mM ClAcK in PBS and incubated at 37°C. At defined time points, 10 µL aliquots of the reaction mixture were withdrawn and quenched by the addition of 1 µL of 0.5 M iodoacetamide. Samples were further diluted with 490 µL of 0.4 M sodium carbonate solution right before RP-HPLC.

RP-HPLC analysis of samples were performed on a 1260 HPLC system (Agilent) equipped with a C18, 4.6 x 250 mm, 10 µm column (Fortis BIO). The mobile phase composed of ddH_2O (A) and methanol (B). The flow rate was 1 mL/min, the injection volume was 20 µL and the temperature of column was maintained at 40°C. Prior to injection, 10 µL of samples were derivatized by mixing with 10 µL of OPA solution for 1 minute in the autosampler. Once samples were injected, the column was washed with 5% B for 5 minutes, followed by a 19-minute gradient elution with 5-100% B. Derivatized amines were detected by FLD using excitation wavelength of 335 nm and emission wavelength of 440 nm.

Molecular dynamic (MD) simulation

For simulation of the αSyn-V52K-SUMO complex, the initial structure was assembled by first predicting a model of SUMO bound to the SIM1 region (^{37}VLYV^{40}) of αSyn using AlphaFold 2.2.0.[2] The structure of the remaining αSyn-V52K sequence was obtained from the NMR structure of αSyn fibrils (PDB: 2N0A) and then concatenated with the AlphaFold-predicted SIM1 region. Since the goal of the simulation is only to show the feasibility of SUMO C-terminus positioning near V52K of αSyn upon SIM1 binding, the initial structure was further manually adjusted to reduce the computational demands of the MD simulation.

MD simulation was performed using GROMACS 2022.3 with OPLS-AA/L force field.[3] The initial structure was solvated with TIP4P water in a dodecahedron box with at least 1.5 nm between the structure and the box edge. To neutralize the system, 150 mM ionic strength was added with appropriate number of counter-ions (Na^+ and Cl^-). Then, the system underwent energy minimization using an Fmax threshold of 1800 kJ/mol and the steepest-descent method. This was followed by two separate equilibration stages: 100

ps NVT (constant number of particles, volume, and temperature) and 100 ps NPT (constant number of particles, pressure, and temperature). Finally, a production run of 100 ns was performed at 300 K using a modified Berendsen thermostat with a time constant of 0.1 ps. Pressure was maintained at 1 bar through the Parrinello-Rahman scheme with a time constant of 2 ps and an isothermal compressibility of 4.5×10^{-5} bar^{-1}. Cutoff distances of 1 nm were used for non-bonded van der Waals and short-range electrostatic interactions, while long-range electrostatic interactions were treated using the particle mesh Ewald (PME) scheme. The LINCS algorithm was employed to constrain bond length.

For simulation of the αSyn-V52K-SUMO-Ubc9 complex, structure of the SUMO-Ubc9 complex was first extracted from the crystal structure (PDB: 1Z5S) and underwent 100 ns MD simulation as described above. The final structure was then modified by attaching αSyn from the initial αSyn-V52K-SUMO complex structure. This attachment was achieved by aligning the SUMO moieties in both structures. The resultant structure served as the starting point for a further 100 ns MD simulation.

Supplementary Figures

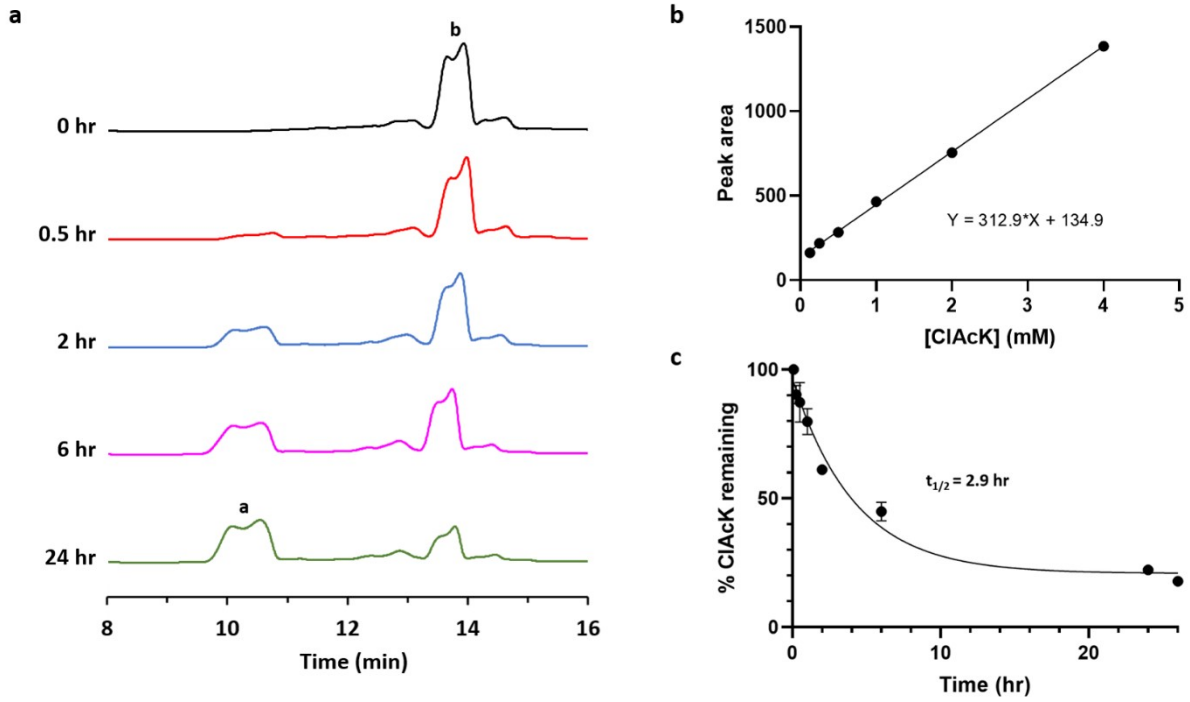


FIGURE S1. Kinetic analysis of 2 mM ClAcK reacting with 10 mM reduced glutathione in phosphate-buffered saline, pH 7.4, at 37°C. (a) Reaction progress of ClAcK and reduced glutathione monitored by HPLC-FLD using excitation wavelength of 335 nm and emission wavelength of 440 nm. Peak a is the reaction product; while b is ClAcK. (b) ClAcK calibration curve. (c) Reaction kinetics of 2 mM ClAcK reacting with 10 mM reduced glutathione. The curves are fitted as pseudo first-order reaction kinetics. Half-life was calculated to be 2.9 hr.

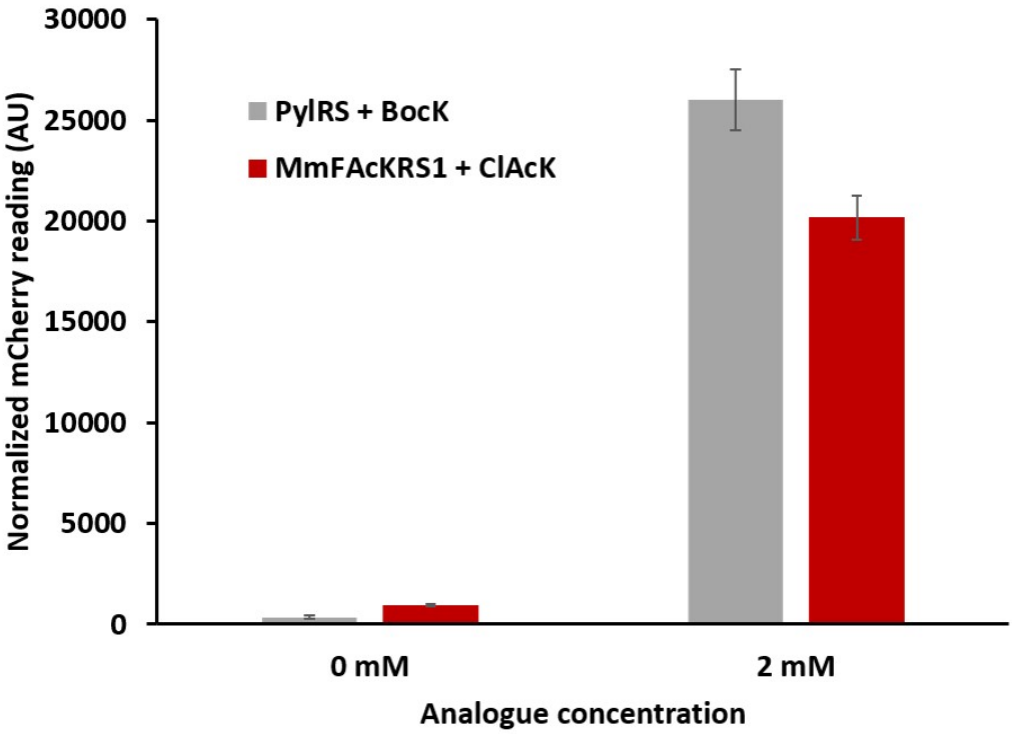
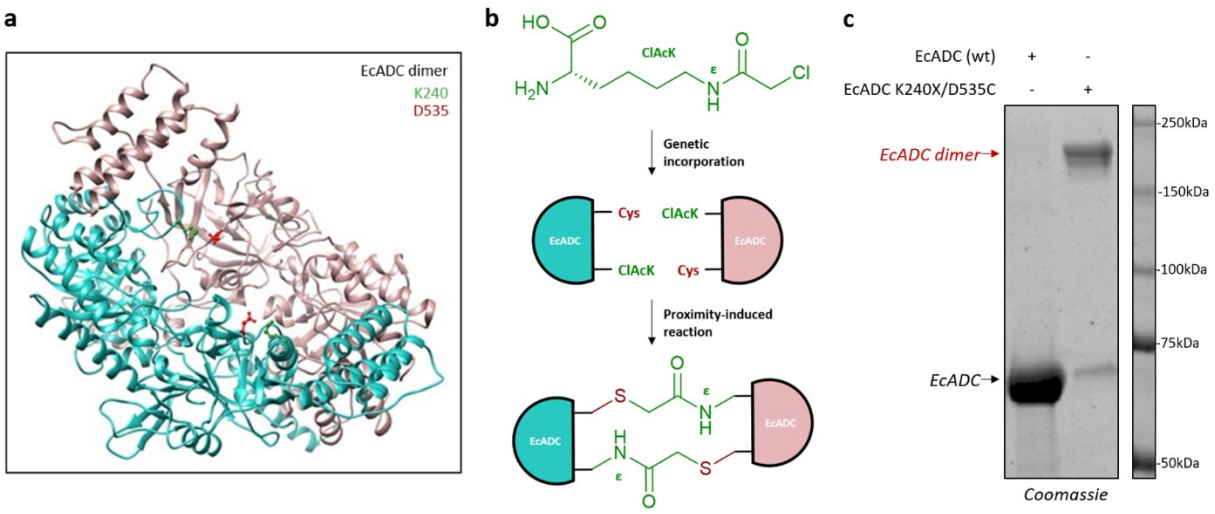


FIGURE S2. Comparison of ClAcK and BocK readthrough efficiencies by MmFAcKRS1 and PylRS using the mCherry reporter assay. Error bars represent SEM based on at least three independent measurements.



d

Protein 1:	arginine decarboxylase, partial [Escherichia coli]		
Accession:	gi|485990858	**Score:**	230.23
Database:	NCBInr	**MW [kDa]:**	70.90
Seq. Coverage [%]:	9.20 %	**pI:**	4.89
		No. of Peptides:	4
Modification(s):	Oxidation		

10	20	30	40	50	60	70	80	90	100	110	120
MSSQEASKML	RTYNIAWWGN	NYYDVNELGH	ISVCPDPDVP	EARVDLAQLV	KTREAQGQRL	PALFCFPQIL	QHRLRSINAA	FKRARESYGY	NGDYFLVYPI	KVNQHRRVIE	SLIHSGEPLG
130	140	150	160	170	180	190	200	210	220	230	240
LEAGSKAELM	AVLAHAGMTR	SVIVCNGYKD	REYIRLALIG	EKMGHKVYLV	IEKMSEIAIV	LDEAERLNVV	PRLGVRARLA	SQGSGKWQSS	GGEKSKFGLA	ATQVLQLVET	LREAGRLDSL
250	260	270	280	290	300	310	320	330	340	350	360
QLLHFHLGSQ	MANIRDIATG	VRESARFYVE	LHKLGVNIQC	FDVGGGLGVD	YEGTRSQSDC	SVNYGLNEYA	NNIIWAIGDA	CEENGLPHPT	VITESGRAVT	AHHTVLVSNI	IGVERNEYTV
370	380	390	400	410	420	430	440	450	460	470	480
PTAPAEDAPR	ALQSMWETWQ	EMHEPGTRRS	LREWLHDSQM	DLHDIHIGYS	SGTFSLQERA	WAEQLYLSMC	HEVQKQLDPQ	NRAHRPIIDE	LQERMADKMY	VNFSLFQSMP	DAWGIDQLFP
490	500	510	520	530	540	550	560	570	580	590	600
VLPLEGLDQV	PERRAVLLDI	TCDSDGAIDH	YIDGDGIATT	MPMPEYDPEN	PPMLGFFMVG	AYQEILGNMH	NLFGDTEAVD	VFVFPDGSVE	VELSDEGDTV	ADMLQYVQLD	PKTLLTQFRD
610	620	630	640								
QVKKTDLDAE	LQQQFLEEFE	AGLYGYTYLE									

Cmpd.	No. of Cmpds.	m/z meas.	Δ m/z [ppm]	z	Rt [min]	Score	Site [%]	P	Range	Sequence	Modification	Type
	1	1491.7291	-4.58	1		25.40		0	174-186	K.MSEIAIVLDEAER.L	Oxidation: 1	CID
	1	1916.1203	25.64	1		99.49		0	338-355	R.AVTAHHTVLVSNIIGVER.N		CID
	1	1630.7946	14.61	1		71.35		0	356-370	R.NEYTVPTAPAEDAPR.A		CID
	1	1476.7768	-10.14	1		33.99		0	443-454	R.AHRPIIDELQER.M		CID

FIGURE S3. Construction of the covalent EcADC dimer. (a) Structure of the EcADC homodimer (PDB ID 3NZQ). Residues K240 and D535 are highlighted in green and red, respectively. Note the head-to-tail arrangement of the homodimer positions K240 of one monomer in close proximity to D535 of the other monomer. (b) Scheme of covalent EcADC dimer formation. (c) SDS-PAGE of purified His-tagged protein from BL21(DE3) co-expressing either EcADC wildtype (wt: 70.7kDa) or EcADC mutant (K240X/D535C: 141.4kDa) with PylT/MmFAcKRS1 in the presence of 2 mM ClAcK. Conjugation yield of EcADC dimer by ClAcK was 82.9%. The identity of the dimer band was further verified by peptide mass fingerprinting. (d) Peptide mass fingerprinting verified the identity of the dimeric band as arginine decarboxylase.

Note: The full gel image can be found in the "Full gels and blots" section of the supporting information.

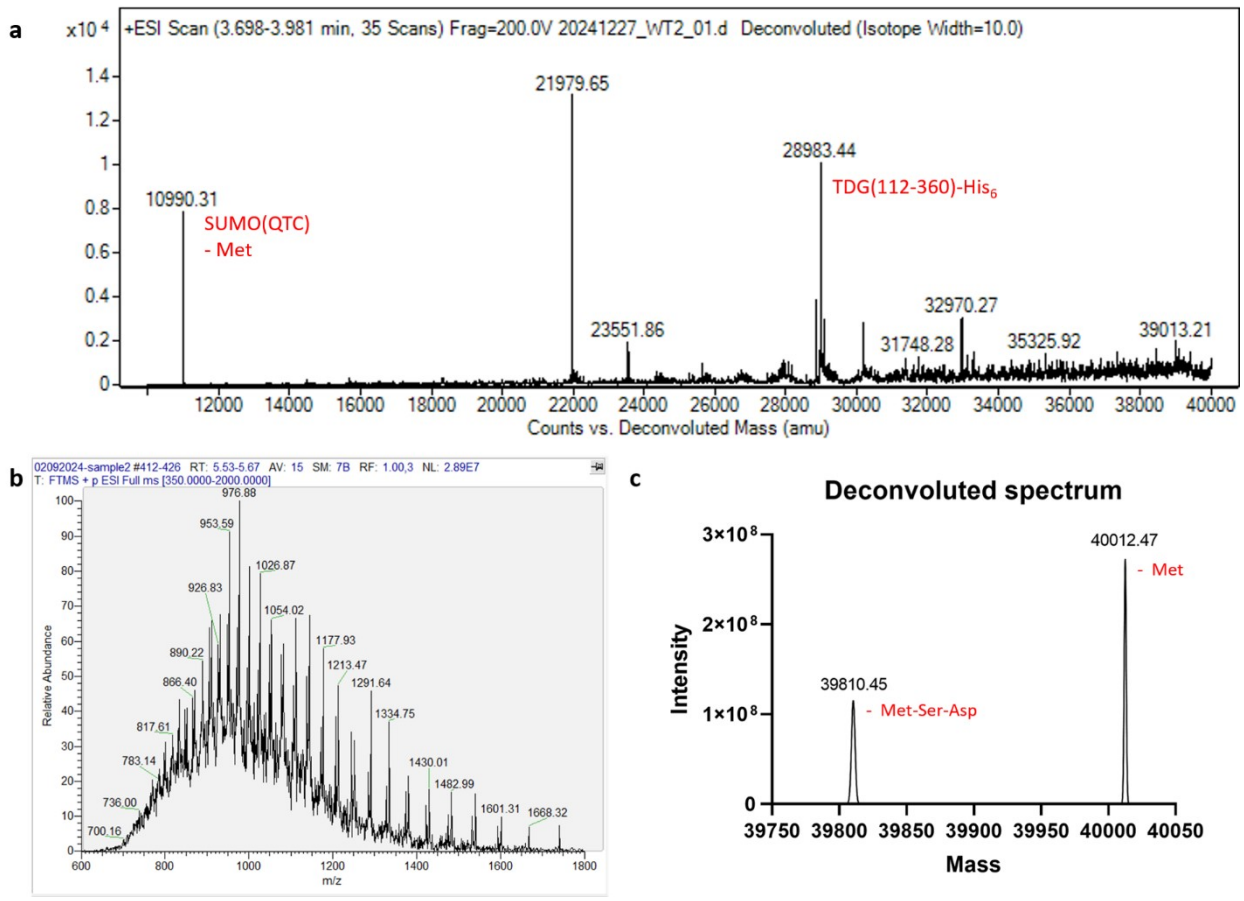


FIGURE S4. Validating the formation of SUMO(QTC)-TDG(112-360)-His_6-K330X conjugate by mass spectroscopy. (a) Deconvoluted spectrum of the co-purified SUMO(QTC) and TDG(112-360)-His_6. The mass of SUMO(QTC) following methionine truncation (theoretical mass = 10,990.37 Da; observed mass = 10,990.31 Da) and the mass of TDG(112-360)-His_6 (theoretical mass = 28,987.06 Da; observed mass = 28,983.44 Da) were identified. Notably, the observed mass for TDG(112-360)-His_6 is -3.62 Da smaller than the theoretical mass, indicating an unknown modification. One possibility is that the two cysteines in the unstructured C-terminal tail (residues 345-360) form two disulfide bonds with other cysteine residues on the protein surface, giving rise to an ~4.03 Da mass reduction. (b) Intact mass spectrum of SUMO(QTC)-TDG(112-360)-K330X-His_6 conjugate. (c) Deconvoluted spectrum of intact SUMO(QTC)-TDG(112-360)-K330X-His_6. The experimental mass of 40,012.47 Da is assigned to the SUMO(QTC)-TDG(112-360)-K330X-His_6 conjugate with two internal disulfide bonds and an N-terminal methionine truncation (theoretical mass = 40,013.43 Da); while the experimental mass of 39,810.45 Da is assigned to the same SUMO(QTC)-TDG(112-360)-K330X-His_6 conjugate with truncation of N-terminal Met-Ser-Asp sequence on SUMO(QTC) (theoretical mass = 39,811.26 Da).

#1	b+	b2+	b3+	Seq.	y+	y2+	y3+	#2
1	130.04987	65.52857	44.02147	E				18
2	243.13393	122.07061	81.71616	L	2904.28351	1452.64540	968.76602	17
3	300.15540	150.58134	100.72332	G	2791.19945	1396.10336	931.07133	16
4	447.19080	224.09904	149.73512	M-Oxidation	2734.17799	1367.59263	912.06418	15
5	576.23339	288.62033	192.74931	E	2587.14259	1294.07493	863.05238	14
6	705.27598	353.14163	235.76351	E	2458.09999	1229.55364	820.03818	13
7	834.31858	417.66293	278.77771	E	2329.05740	1165.03234	777.02398	12
8	949.34552	475.17640	317.12002	D	2200.01481	1100.51104	734.00979	11
9	1048.41393	524.71060	350.14283	V	2084.98787	1042.99757	695.66747	10
10	1161.49800	581.25264	387.83752	I	1985.91945	993.46336	662.64467	9
11	1290.54059	645.77393	430.85171	E	1872.83539	936.92133	624.94998	8
12	1389.60900	695.30814	463.87452	V	1743.79279	872.40004	581.93578	7
13	1552.67233	776.83980	518.22896	Y	1644.72438	822.86583	548.91298	6
14	1680.73091	840.86909	560.91515	Q	1481.66105	741.33416	494.55854	5
15	1809.77350	905.39039	603.92935	E	1353.60247	677.30488	451.87234	4
16	1937.83208	969.41968	646.61554	Q	1224.55988	612.78358	408.85814	3
17	2038.87976	1019.94352	680.29810	T	1096.50130	548.75429	366.17195	2
18				C-TDGK33...	995.45363	498.23045	332.48939	1

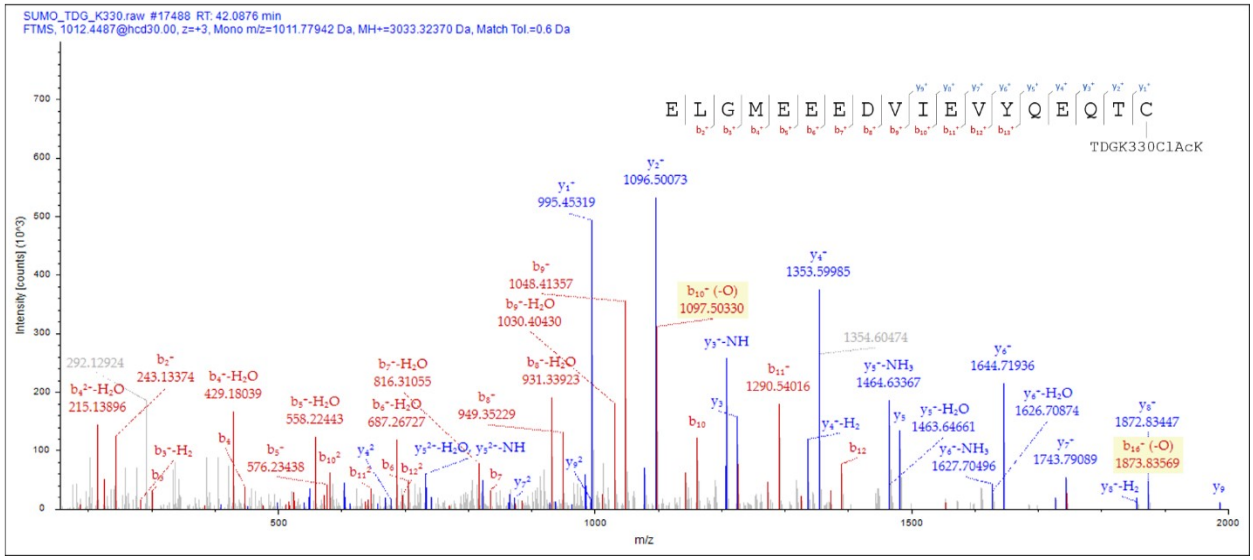


FIGURE S5. MS/MS spectrum identifying the crosslink between SUMO(QTC) C-terminus cysteine and the ClAcK modification at position K330 of TDG(112-360). The fragmentation pattern corresponds to the C-terminal tryptic peptide from SUMO(QTC) conjugated to the 7-residue ClAcK-bearing tryptic fragment of TDGK330X: MAVXEEK; Mass shift associated with the formation of thioether linkage between SUMO C-terminus cysteine and this fragment: +873.4266

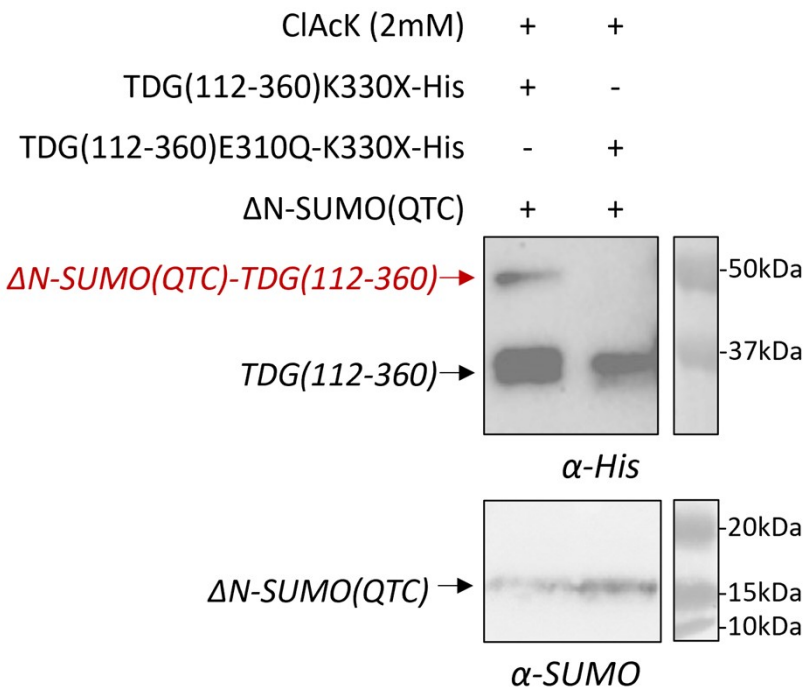


FIGURE S6. Verification of ΔN-SUMO(QTC) expression. Cell lysates of BL21(DE3) co-expressing different TDG(112-360)-His constructs and ΔN-SUMO(QTC) were analyzed by Western blotting using either anti-His or anti-SUMO antibodies. ΔN-SUMO(QTC) was expressed in both constructs, but the conjugation yields varied between TDG(112-360) and its SIM-defective mutant (E310Q).

Note: The full gel image can be found in the "Full gels and blots" section of the supporting information.

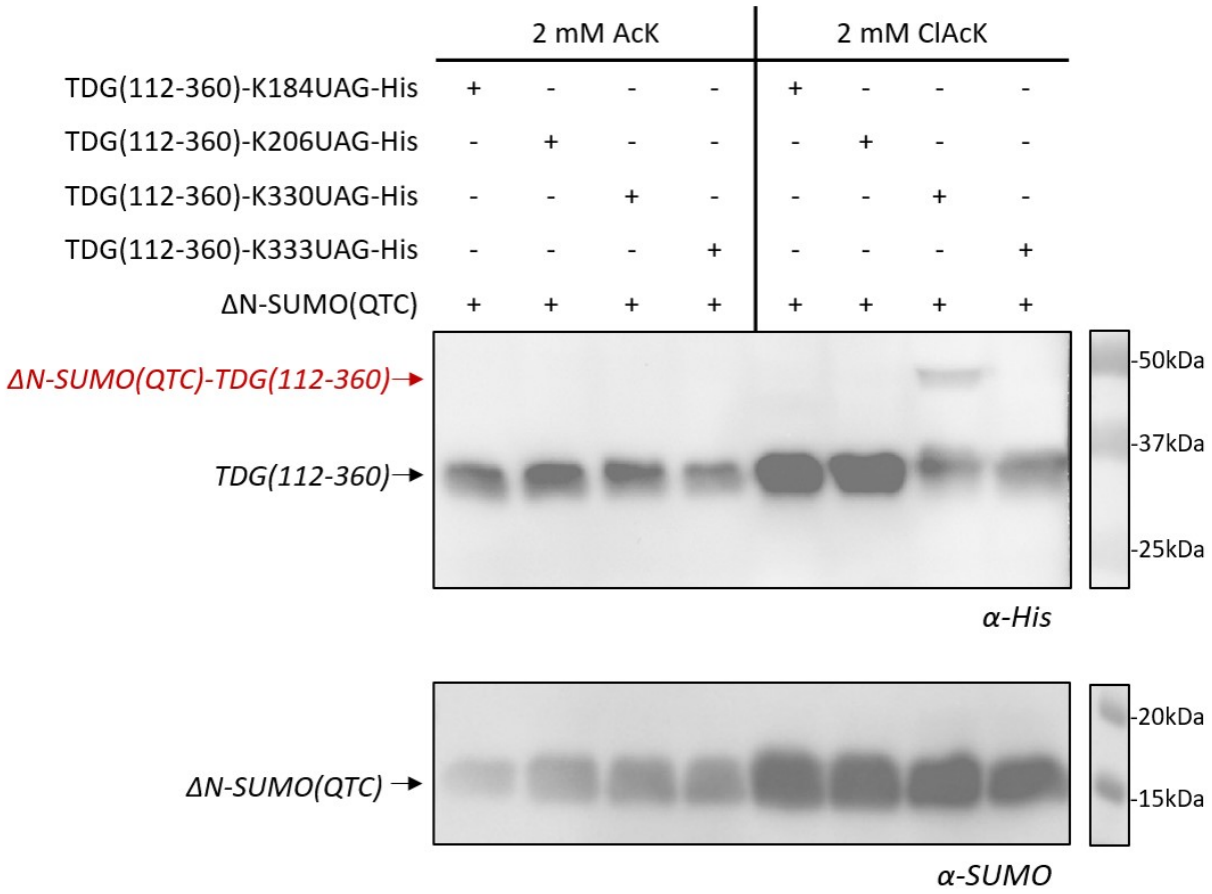


FIGURE S7. Specific conjugation of ΔN-SUMO(QTC) with ClAcK installed at position 330 of TDG(112-360)-His. Different TDG(112-360)-His constructs were co-expressed with ΔN-SUMO(QTC) in the presence of ClAcK or AcK (Nε-acetyl-lysine). Western blot analysis of cell lysates, using either anti-His or anti-SUMO antibodies, was performed to confirm the expression of ΔN-SUMO(QTC) and to evaluate the chemical and spatial specificity of the conjugation.

Note: The full gel image can be found in the "Full gels and blots" section of the supporting information.

#1	b^+	b^{2+}	b^{3+}	Seq.	y^+	y^{2+}	y^{3+}	#2
1	88.03930	44.52329	30.01795	S				27
2	217.08190	109.04459	73.03215	E	3077.40108	1539.20418	1026.47188	26
3	316.15031	158.57879	106.05496	V	2948.35848	1474.68288	983.45768	25
4	444.20889	222.60808	148.74115	Q	2849.29007	1425.14867	950.43487	24
5	557.29295	279.15011	186.43584	L	2721.23149	1361.11938	907.74868	23
6	671.33588	336.17158	224.45014	N	2608.14743	1304.57735	870.05399	22
7	808.39479	404.70103	270.13645	H	2494.10450	1247.55589	832.03968	21
8	939.43528	470.22128	313.81661	M	2357.04559	1179.02643	786.35338	20
9	1054.46222	527.73475	352.15892	D	2226.00510	1113.50619	742.67322	19
10	1169.48916	585.24822	390.50124	D	2110.97816	1055.99272	704.33090	18
11	1306.54807	653.77768	436.18754	H	1995.95122	998.47925	665.98859	17
12	1407.59575	704.30151	469.87010	T	1858.89231	929.94979	620.30229	16
13	1520.67982	760.84355	507.56479	L	1757.84463	879.42595	586.61973	15
14	1617.73258	809.36993	539.91571	P	1644.76056	822.88392	548.92504	14
15	1674.75404	837.88066	558.92287	G	1547.70780	774.35754	516.57412	13
16	1878.82060	939.91394	626.94505	K-Chloroac...	1490.68634	745.84681	497.56696	12
17	2041.88393	1021.44560	681.29949	Y	1286.61978	643.81353	429.54478	11
18	2098.90539	1049.95633	700.30665	G	1123.55645	562.28187	375.19034	10
19	2211.98945	1106.49837	738.00134	I	1066.53499	533.77113	356.18318	9
20	2269.01092	1135.00910	757.00849	G	953.45093	477.22910	318.48849	8
21	2416.07933	1208.54330	806.03129	F	896.42946	448.71837	299.48134	7
22	2517.12701	1259.06714	839.71385	T	749.36105	375.18416	250.45853	6
23	2631.16994	1316.08861	877.72816	N	648.31337	324.66032	216.77597	5
24	2762.21042	1381.60885	921.40832	M	534.27044	267.63886	178.76167	4
25	2861.27884	1431.14306	954.43113	V	403.22996	202.11862	135.08150	3
26	2990.32143	1495.66435	997.44533	E	304.16155	152.58441	102.05870	2
27				R	175.11895	88.06311	59.04450	1

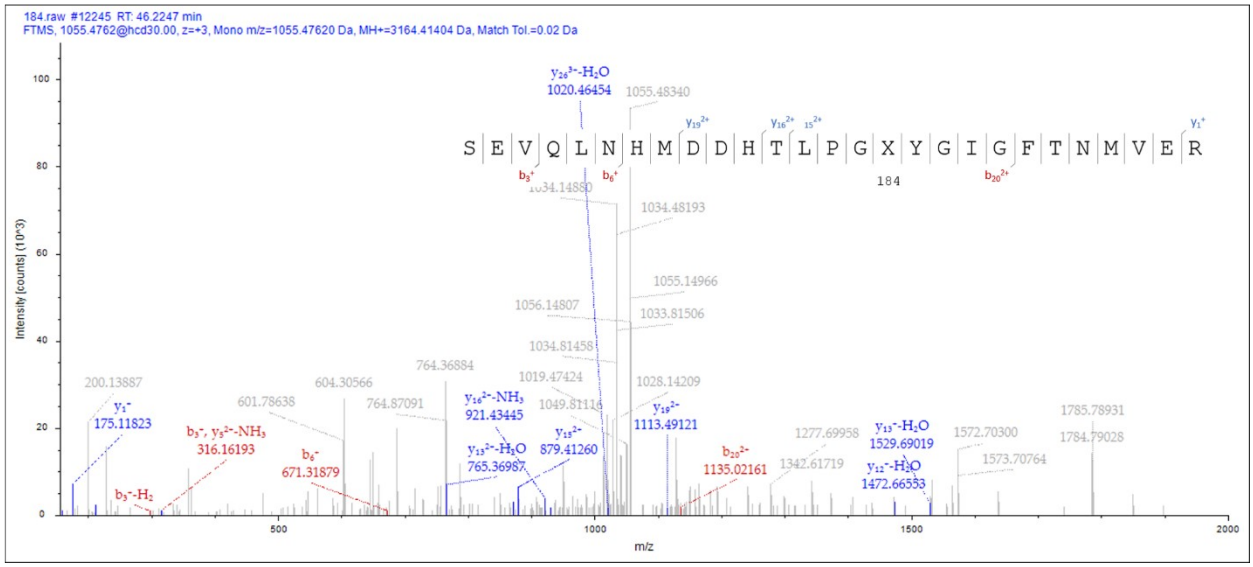


FIGURE S8. An MS/MS spectrum confirming the incorporation of ClAcK (denoted as X) at position 184 of TDG(112-360).

#1	b+	b2+	Seq.	y+	y2+	#2
1	116.03422	58.52075	D			8
2	229.11828	115.06278	L	942.44463	471.72595	7
3	316.15031	158.57879	S	829.36057	415.18392	6
4	403.18234	202.09481	S	742.32854	371.66791	5
5	607.24889	304.12808	K-Chloroac...	655.29651	328.15189	4
6	736.29149	368.64938	E	451.22996	226.11862	3
7	883.35990	442.18359	F	322.18737	161.59732	2
8			R	175.11895	88.06311	1

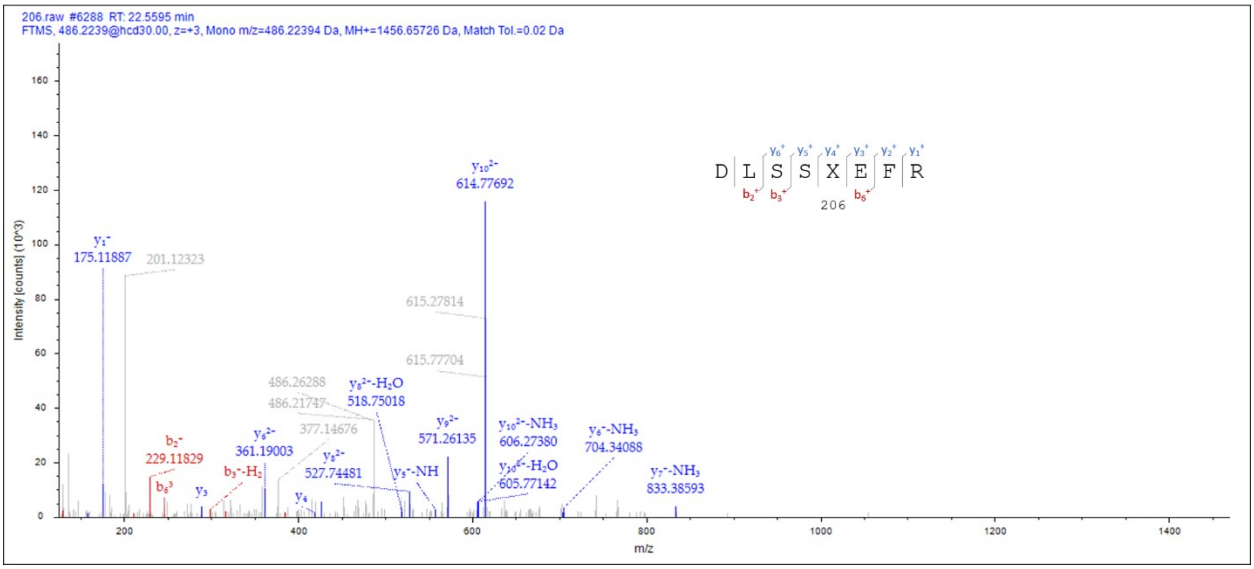


FIGURE S9. An MS/MS spectrum confirming the incorporation of ClAcK (denoted as X) at position 206 of TDG(112-360).

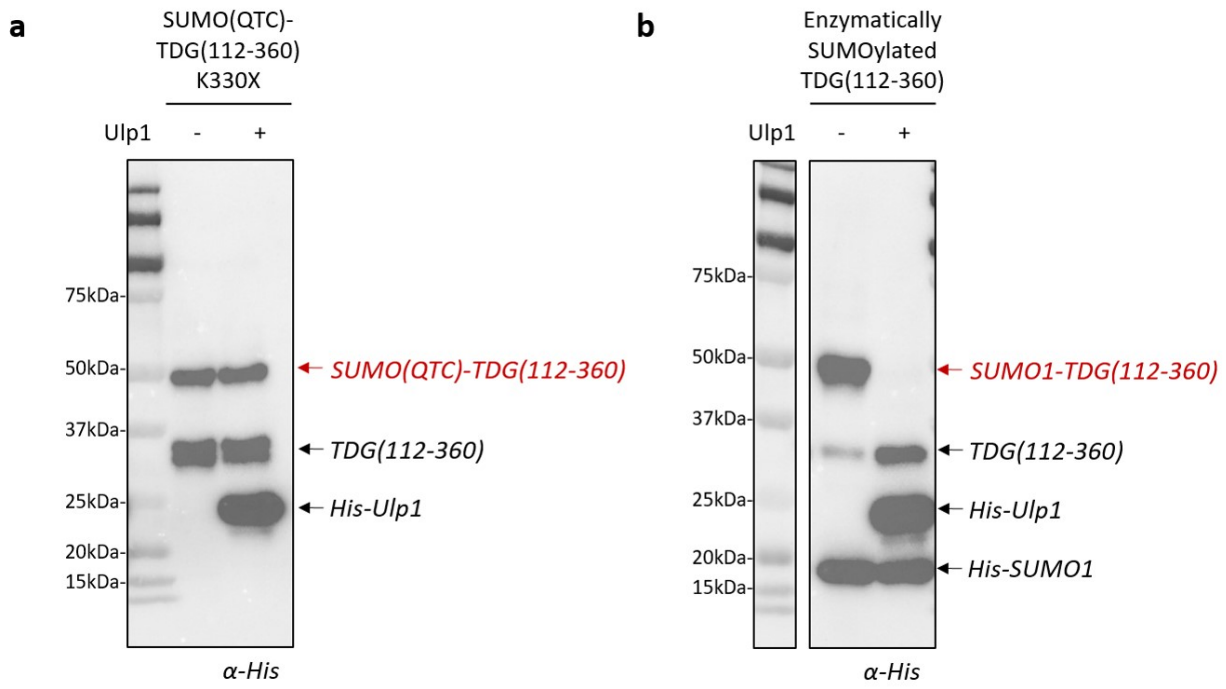


FIGURE S10. Resistance of ClAcK-mediated SUMO-TDG conjugate to Ulp1 SUMO protease cleavage. Western blot analysis of (a) ClAcK-mediated SUMO-TDG conjugate and (b) enzymatically SUMOylated TDG(112-360) following incubation with 0.4 mg/mL Ulp1 overnight at 30°C.

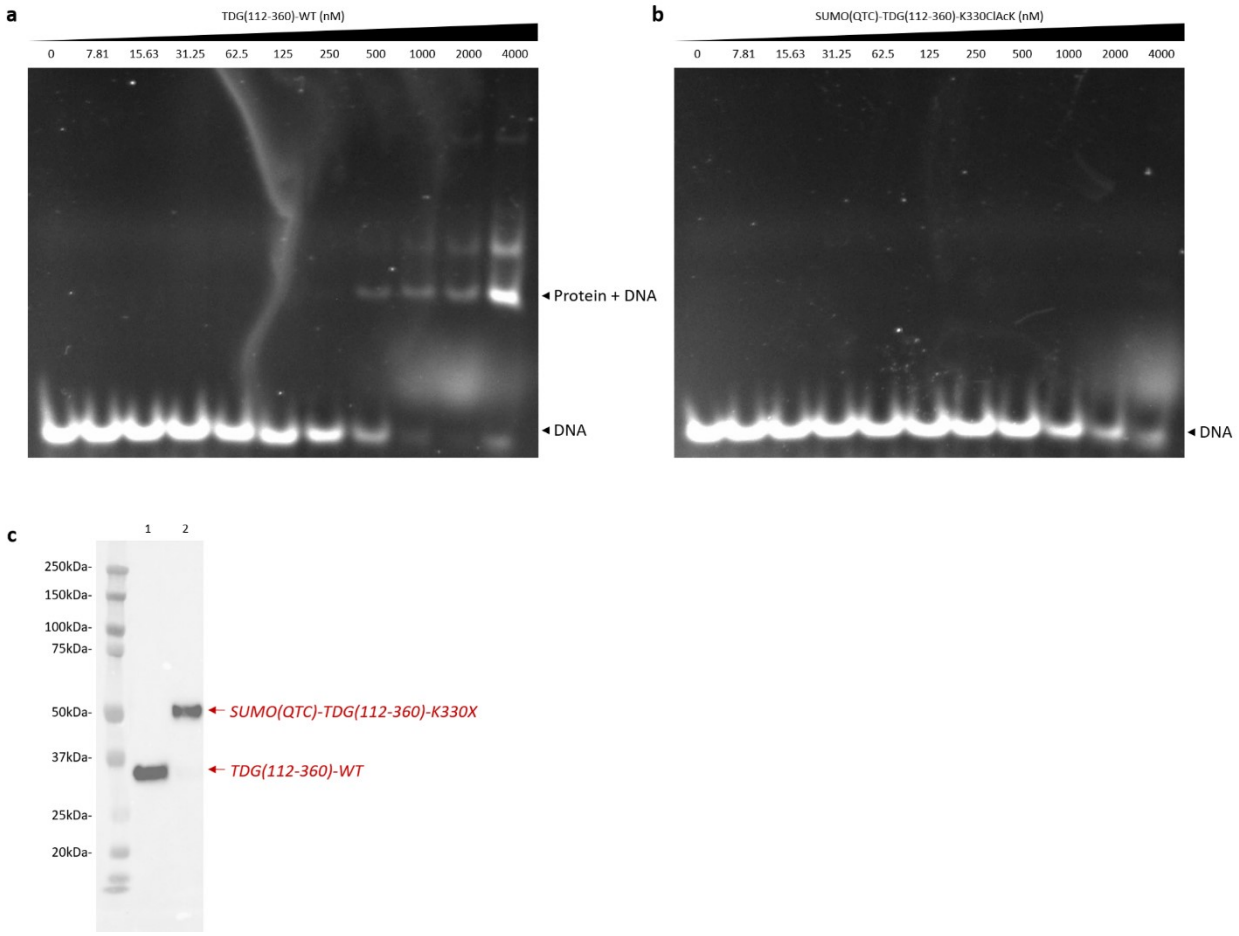


FIGURE S11. Effect of ClAcK-mediated SUMOylation of TDG(112-360)-K330X on its binding to G·U mismatch-containing DNA. Electrophoretic mobility shift assays (EMSA) were performed to assess the equilibrium binding of (a) TDG(112-360) or (b) SUMO(QTC)-TDG(112-360)-K330X to DNA substrate (10 nM) containing a G·U mismatch on a 6% native polyacrylamide gel. (c) Western blot analysis of 8 μM purified TDG(112-360)-WT and 8 μM purified SUMO(QTC)-TDG(112-360)-K330X used for EMSA.

#1	b+	b2+	b3+	Seq.	y+	y2+	y3+	#2
1	130.04987	65.52857	44.02147	E				18
2	243.13393	122.07061	81.71616	L	3676.72441	1838.86585	1226.24632	17
3	300.15540	150.58134	100.72332	G	3563.64035	1782.32381	1188.55163	16
4	447.19080	224.09904	149.73512	M-Oxidation	3506.61889	1753.81308	1169.54448	15
5	576.23339	288.62033	192.74931	E	3359.58349	1680.29538	1120.53268	14
6	705.27598	353.14163	235.76351	E	3230.54089	1615.77409	1077.51848	13
7	834.31858	417.66293	278.77771	E	3101.49830	1551.25279	1034.50428	12
8	949.34552	475.17640	317.12002	D	2972.45571	1486.73149	991.49009	11
9	1048.41393	524.71060	350.14283	V	2857.42877	1429.21802	953.14777	10
10	1161.49800	581.25264	387.83752	I	2758.36035	1379.68381	920.12497	9
11	1290.54059	645.77393	430.85171	E	2645.27629	1323.14178	882.43028	8
12	1389.60900	695.30814	463.87452	V	2516.23369	1258.62049	839.41608	7
13	1552.67233	776.83980	518.22896	Y	2417.16528	1209.08628	806.39328	6
14	1680.73091	840.86909	560.91515	Q	2254.10195	1127.55461	752.03884	5
15	1809.77350	905.39039	603.92935	E	2126.04337	1063.52533	709.35264	4
16	1937.83208	969.41968	646.61554	Q	1997.00078	999.00403	666.33844	3
17	2038.87976	1019.94352	680.29810	T	1868.94220	934.97474	623.65225	2
18				C-aSynK9...	1767.89453	884.45090	589.96969	1

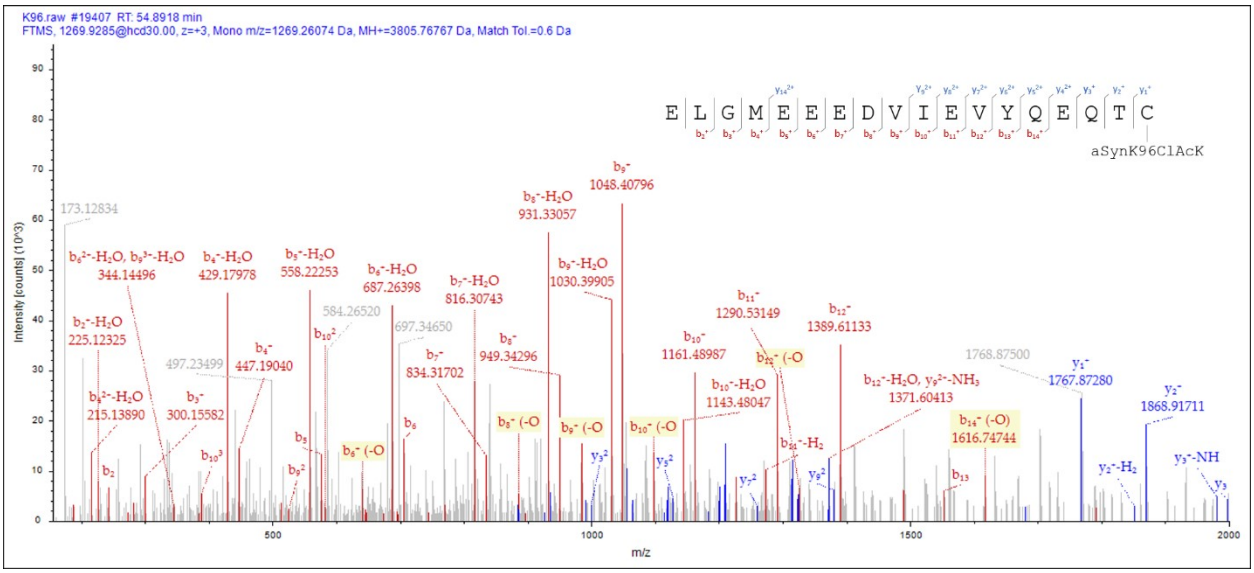


FIGURE S12. MS/MS spectrum identifying the crosslink between SUMO(QTC) C-terminus cysteine and the ClAcK modification at position K96 of αSyn. The fragmentation pattern corresponds to the C-terminal tryptic peptide from SUMO(QTC) conjugated to the ClAcK-bearing tryptic fragment of αSynK96X: TVEGAGSIAA ATGFVXK; Mass shift associated with the formation of thioether linkage between SUMO C-terminus cysteine and this fragment: +1645.8675

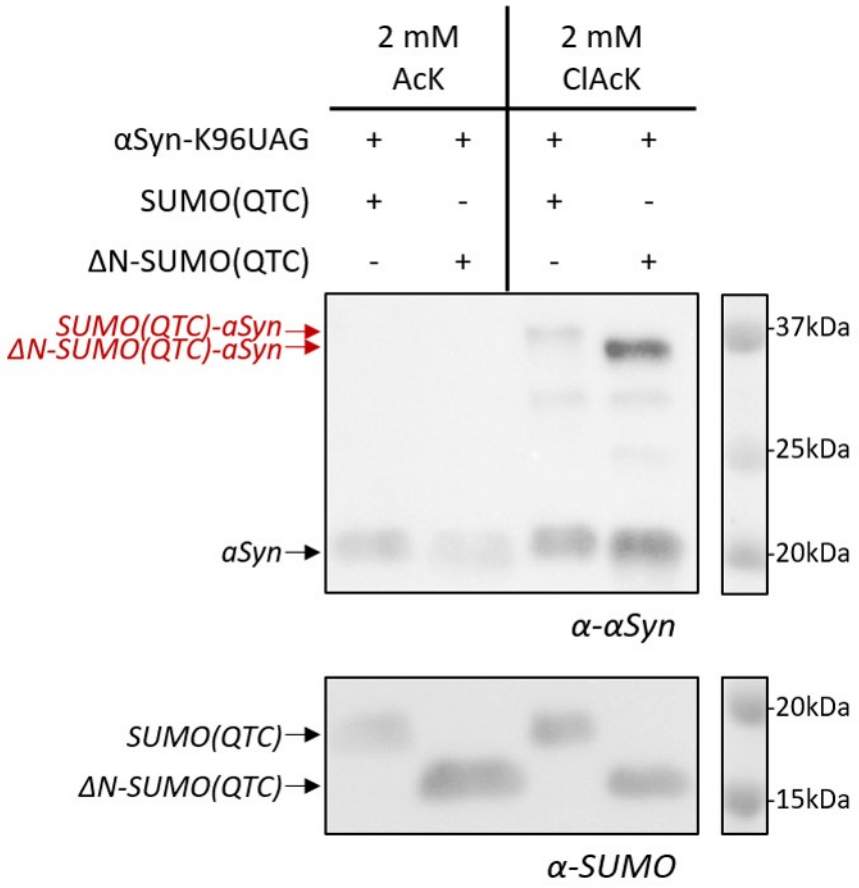


FIGURE S13. Comparison of conjugation rate between SUMO(QTC) and ΔN-SUMO(QTC) to αSyn-His. Boiled cell lysates of BL21(DE3) co-expressing different αSyn-K96UAG and SUMO(QTC) variants were analyzed by Western blotting using either anti-αSyn or anti-SUMO antibodies. SUMO(QTC) variants were expressed in all constructs, but only formed conjugates with αSyn when ClAcK, and not AcK (Nε-acetyl-lysine), was incorporated at position 96 of αSyn. Conjugation yield determined by ImageJ: SUMO(QTC), 22.5%; ΔN-SUMO(QTC), 48.5%

Note: The full gel image can be found in the "Full gels and blots" section of the supporting information.

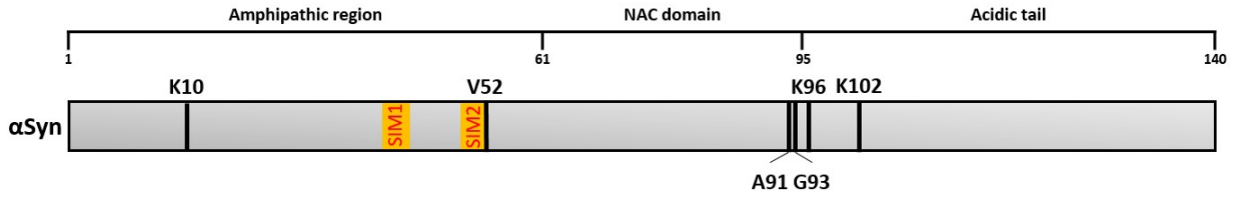


FIGURE S14. Illustration of the positions of the alternate ClAcK-incorporation sites chosen on αSyn. The previously identified SIM sites,[4] SIM1 and SIM2, are highlighted in gold.

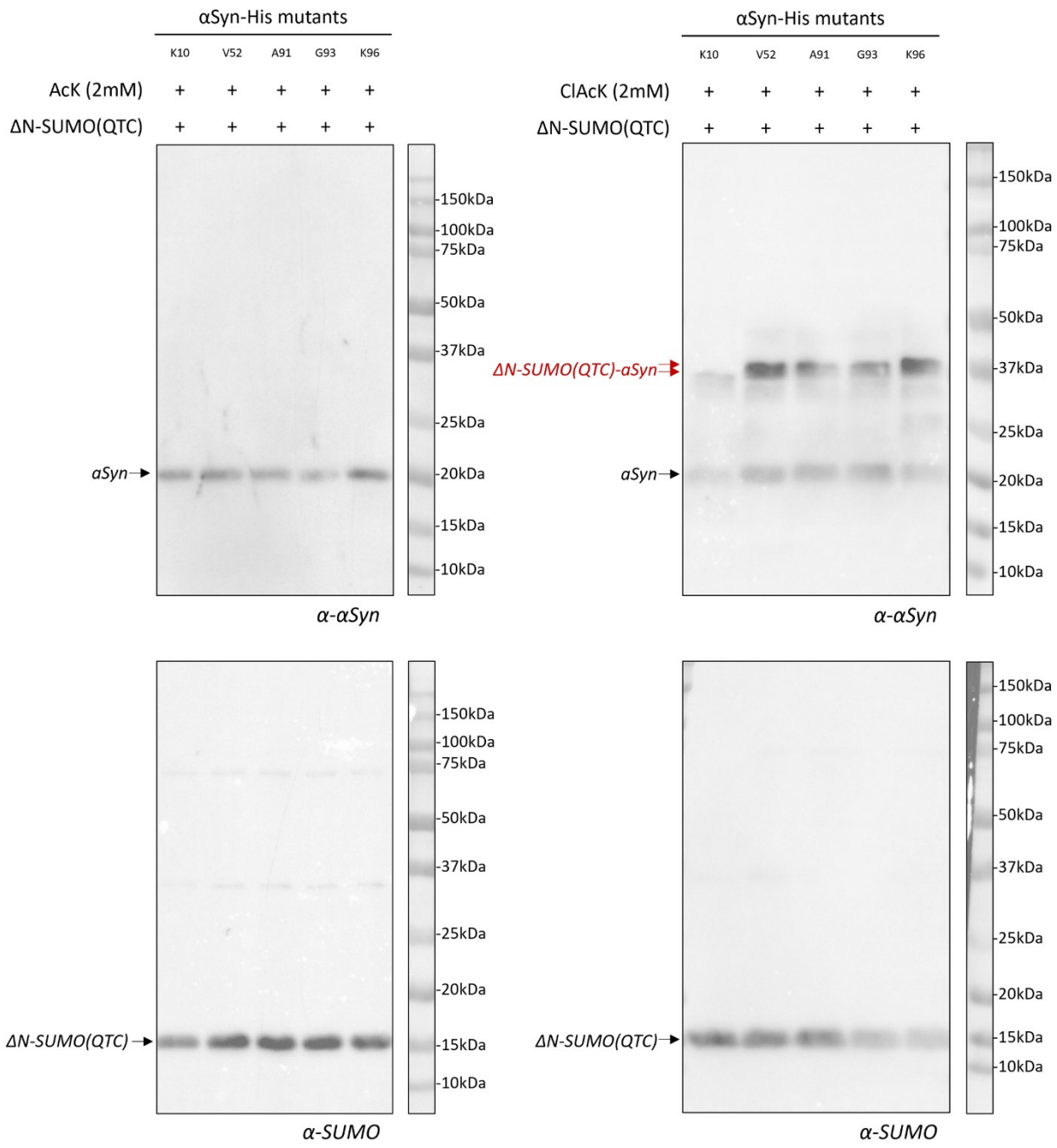


FIGURE S15. ClAcK-mediated SUMOylation of αSyn at alternate sites. Boiled cell lysates of BL21(DE3) co-expressing different αSyn constructs and ΔN-SUMO(QTC) were analyzed by Western blotting using either anti-αSyn or anti-SUMO antibodies. ΔN-SUMO(QTC) was expressed in all constructs and formed conjugates with αSyn only when ClAcK, but not AcK (Nε-acetyl-lysine), was incorporated into αSyn. Regarding the anomalous migration pattern of ΔN-SUMO(QTC)- αSyn-K10X conjugate, previous study has observed that αSyn carrying the same modification, but at different attachment sites, can exhibit distinct migration patterns on SDS-PAGE.[5] Conjugation yield determined by ImageJ: αSyn-K10X, 40.9%; αSyn-V52X, 61%; αSyn-A91X, 55.1%; αSyn-G93X, 50%; αSyn-K96X, 68%

Note: The full gel image can be found in the "Full gels and blots" section of the supporting information.

#1	b+	b2+	b3+	Seq.	y+	y2+	y3+	#2
1	130.04987	65.52857	44.02147	E				18
2	243.13393	122.07061	81.71616	L	2673.22691	1337.11710	891.74716	17
3	300.15540	150.58134	100.72332	G	2560.14285	1280.57506	854.05247	16
4	447.19080	224.09904	149.73512	M-Oxidation	2503.12139	1252.06433	835.04531	15
5	576.23339	288.62033	192.74931	E	2356.08599	1178.54663	786.03351	14
6	705.27598	353.14163	235.76351	E	2227.04339	1114.02534	743.01932	13
7	834.31858	417.66293	278.77771	E	2098.00080	1049.50404	700.00512	12
8	949.34552	475.17640	317.12002	D	1968.95821	984.98274	656.99092	11
9	1048.41393	524.71060	350.14283	V	1853.93127	927.46927	618.64861	10
10	1161.49800	581.25264	387.83752	I	1754.86285	877.93506	585.62580	9
11	1290.54059	645.77393	430.85171	E	1641.77879	821.39303	547.93111	8
12	1389.60900	695.30814	463.87452	V	1512.73619	756.87174	504.91692	7
13	1552.67233	776.83980	518.22896	Y	1413.66778	707.33753	471.89411	6
14	1680.73091	840.86909	560.91515	Q	1250.60445	625.80586	417.53967	5
15	1809.77350	905.39039	603.92935	E	1122.54587	561.77658	374.85348	4
16	1937.83208	969.41968	646.61554	Q	993.50328	497.25528	331.83928	3
17	2038.87976	1019.94352	680.29810	T	865.44470	433.22599	289.15309	2
18				C-aSynK1...	764.39703	382.70215	255.47053	1

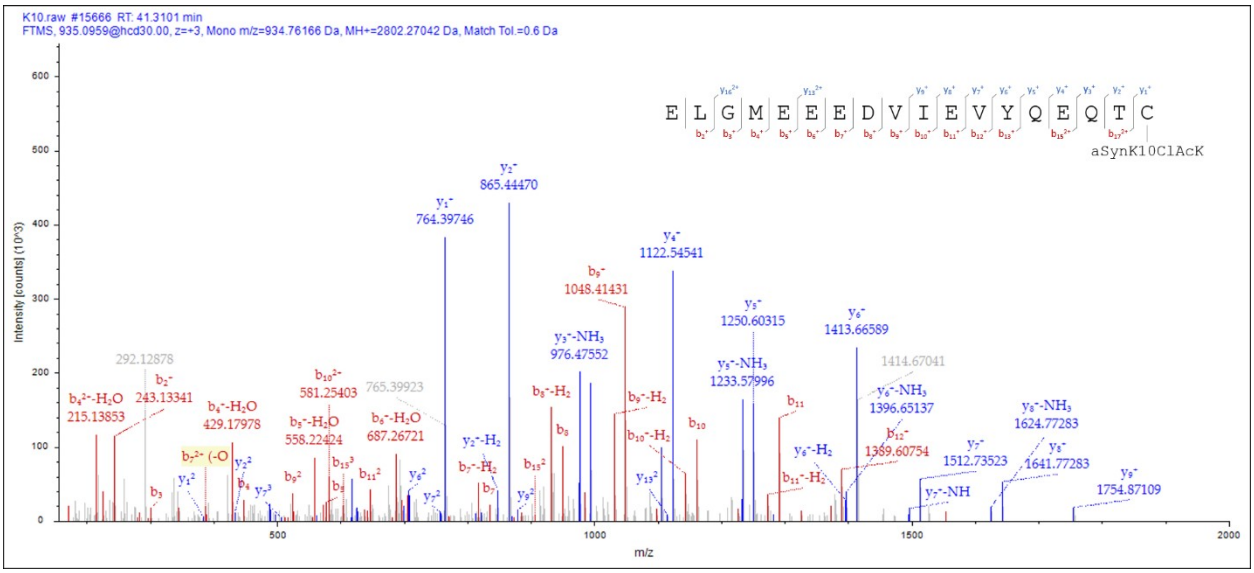


FIGURE S16. MS/MS spectrum identifying the crosslink between SUMO(QTC) C-terminus cysteine and the ClAcK modification at position K10 of αSyn. The fragmentation pattern corresponds to the C-terminal tryptic peptide from SUMO(QTC) conjugated to the ClAcK-bearing tryptic fragment of αSynK10X: GLSXAK; Mass shift associated with the formation of thioether linkage between SUMO C-terminus cysteine and this fragment: +642.3700

#1	b+	b2+	b3+	b4+	Seq.	y+	y2+	y3+	y4+	#2
1	130.04987	65.52857	44.02147	33.26792	E					18
2	243.13393	122.07061	81.71616	61.53894	L	3394.56641	1697.78685	1132.19366	849.39706	17
3	300.15540	150.58134	100.72332	75.79431	G	3281.48235	1641.24481	1094.49897	821.12605	16
4	447.19080	224.09904	149.73512	112.55316	M-Oxidation	3224.46089	1612.73408	1075.49181	806.87068	15
5	576.23339	288.62033	192.74931	144.81380	E	3077.42549	1539.21638	1026.48001	770.11183	14
6	705.27598	353.14163	235.76351	177.07445	E	2948.38289	1474.69509	983.46582	737.85118	13
7	834.31858	417.66293	278.77771	209.33510	E	2819.34030	1410.17379	940.45162	705.59053	12
8	949.34552	475.17640	317.12002	238.09184	D	2690.29771	1345.65249	897.43742	673.32988	11
9	1048.41393	524.71060	350.14283	262.85894	V	2575.27077	1288.13902	859.09511	644.57315	10
10	1161.49800	581.25264	387.83752	291.12996	I	2476.20235	1238.60481	826.07230	619.80605	9
11	1290.54059	645.77393	430.85171	323.39060	E	2363.11829	1182.06278	788.37761	591.53503	8
12	1389.60900	695.30814	463.87452	348.15771	V	2234.07569	1117.54149	745.36342	559.27438	7
13	1552.67233	776.83980	518.22896	388.92354	Y	2135.00728	1068.00728	712.34061	534.50728	6
14	1680.73091	840.86909	560.91515	420.93818	Q	1971.94395	986.47561	657.98617	493.74145	5
15	1809.77350	905.39039	603.92935	453.19883	E	1843.88537	922.44633	615.29998	461.72680	4
16	1937.83208	969.41968	646.61554	485.21348	Q	1714.84278	857.92503	572.28578	429.46615	3
17	2038.87976	1019.94352	680.29810	510.47540	T	1586.78420	793.89574	529.59959	397.45151	2
18					C-aSynV5...	1485.73653	743.37190	495.91703	372.18959	1

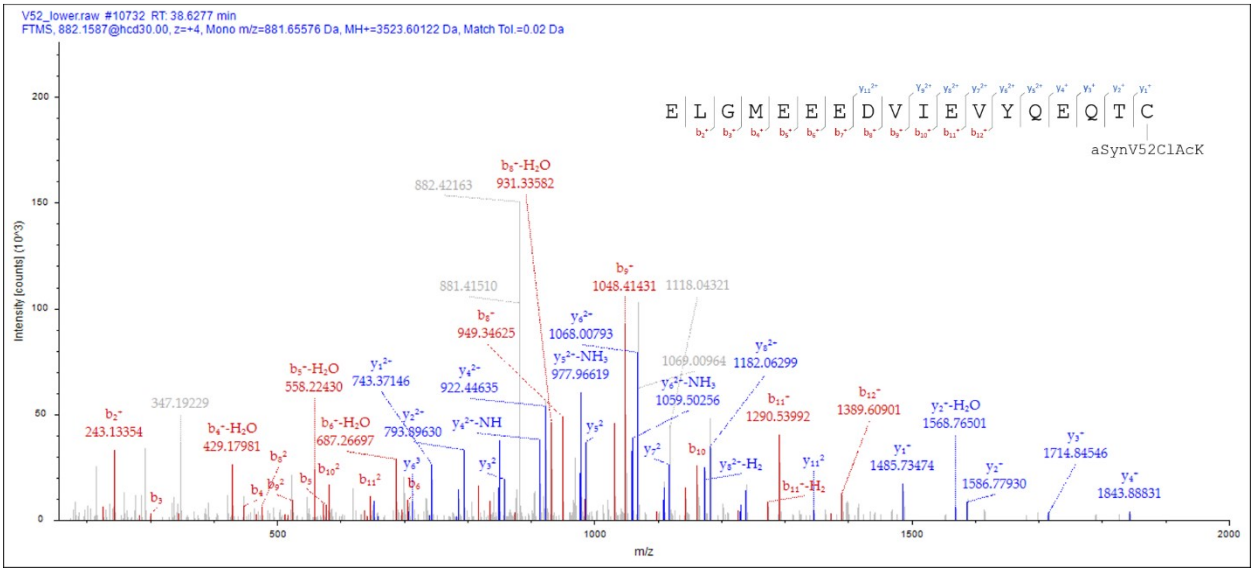


FIGURE S17. MS/MS spectrum identifying the crosslink between SUMO(QTC) C-terminus cysteine and the ClAcK modification at position V52 of αSyn. The fragmentation pattern corresponds to the C-terminal tryptic peptide from SUMO(QTC) conjugated to the ClAcK-bearing tryptic fragment of αSynV52X: EGVVHGXATV AEK; Mass shift associated with the formation of thioether linkage between SUMO C-terminus cysteine and this fragment: +1363.7095

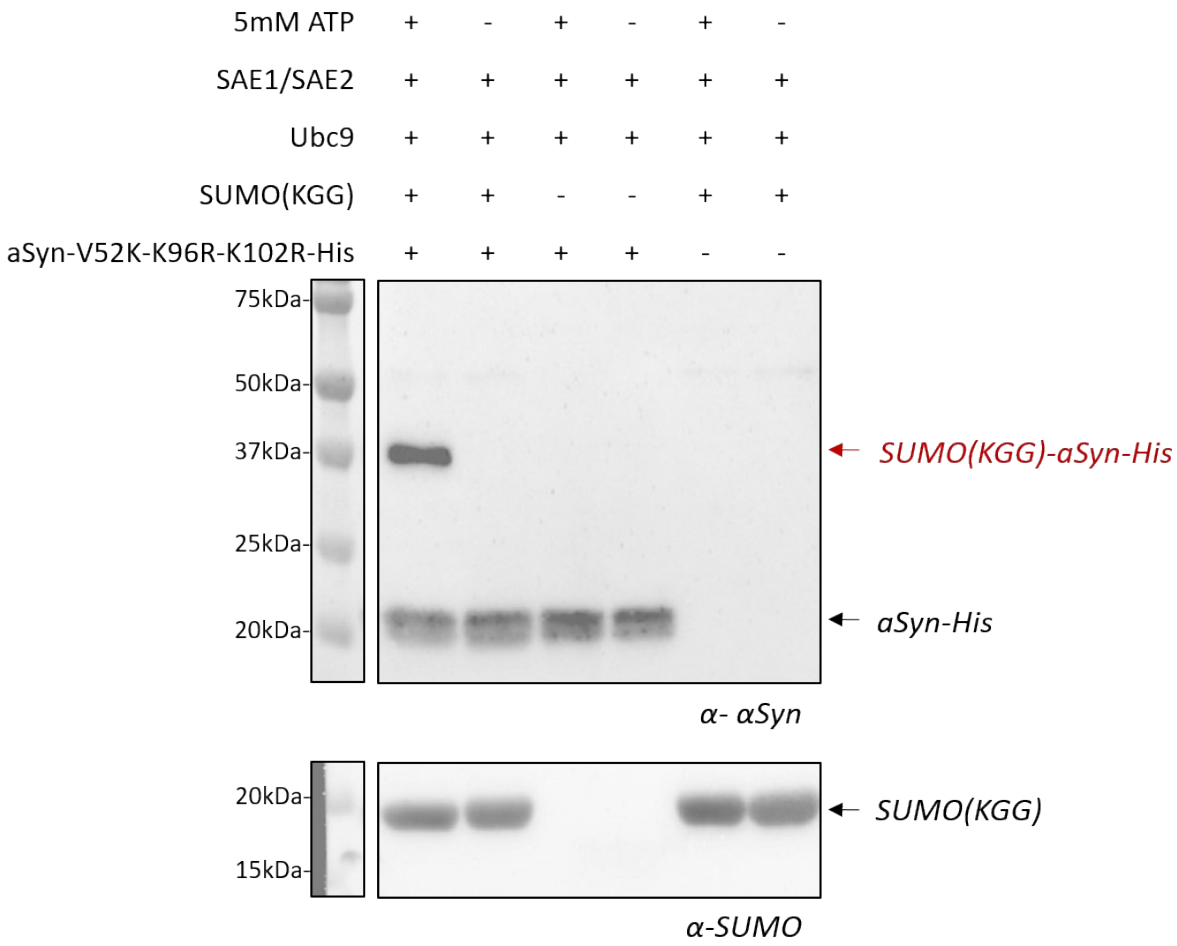


FIGURE S18. *In vitro* enzymatic SUMOylation. Purified αSyn-V52K-K96R-K102R was incubated at 37°C for 2 hours with a reaction mixture containing SUMO-E1 activating enzyme (SAE1/SAE2), SUMO-E2 conjugating enzyme (Ubc9), ATP, and SUMO(KGG), a SUMO variant with the C-terminal TGG-motif mutated to KGG. The products were analyzed by western blotting with anti-αSyn and anti-SUMO antibodies. Note that the full gel image can be found in the "Full gels and blots" section of the supporting information.

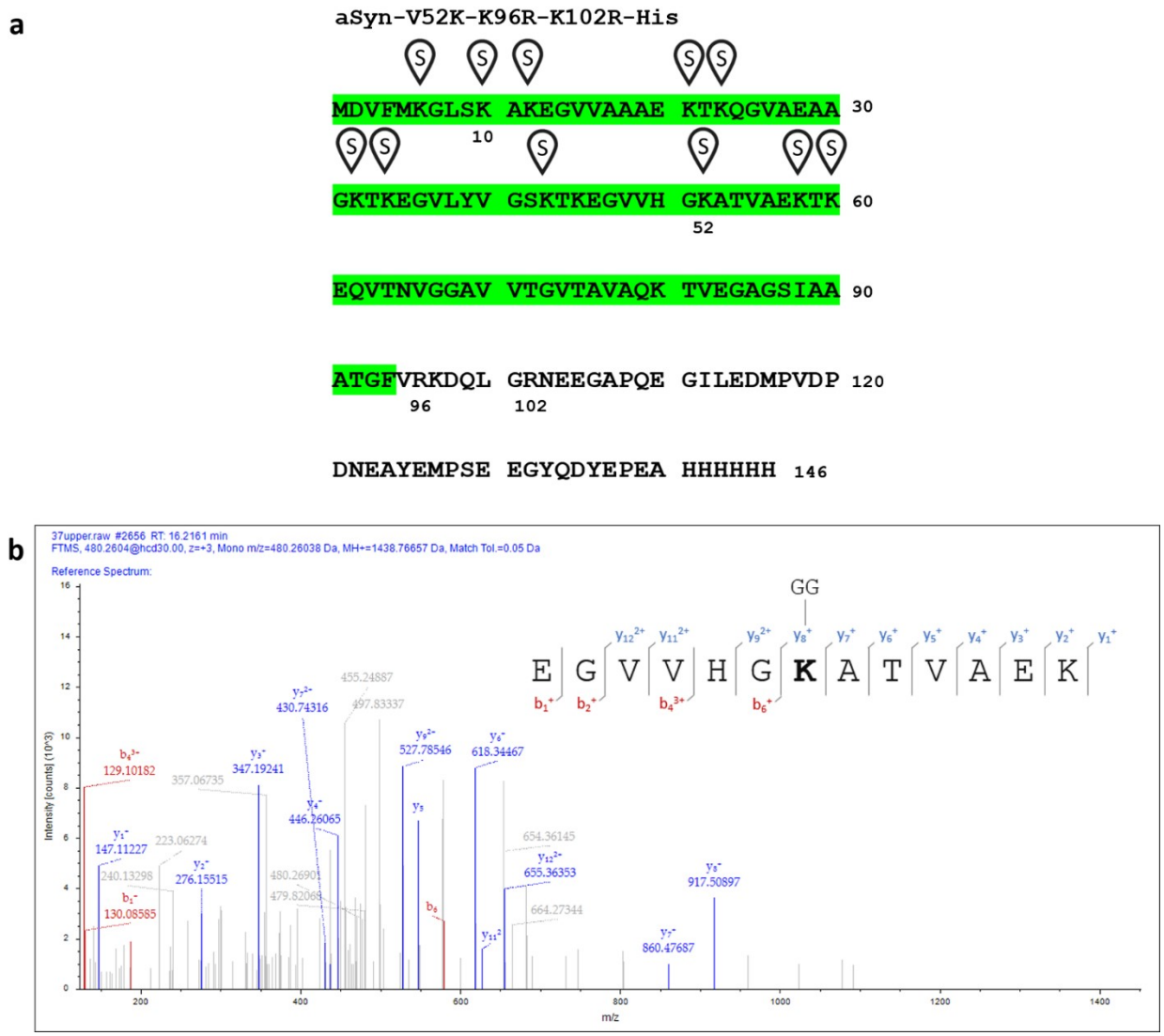


FIGURE S19. LC-MS/MS analysis of αSyn-V52K-K96R-K102R enzymatically SUMOylated using SUMO(KGG). (a) Sequence coverage of the trypsin-digested, SUMOylated αSyn-V52K-K96R-K102R in LC-MS/MS analysis. Lysine residues identified as SUMOylated based on the presence of the di-glycine signature are annotated with "S". (b) Representative MS/MS spectrum of the αSyn-V52K-K96R-K102R tryptic fragment showing SUMOylation at position V52K. The relevant lysine residue (K52) exhibiting the di-glycine signature is highlighted in bold.

Results of MD simulation

MD simulation of αSyn-V52K with SUMO bound to SIM1 (^{37}VLYV^{40}) was performed as detailed in the experimental methods. The trajectory was subsequently analyzed using VMD software. Our analysis showed that in 25% of the conformations sampled, the distance between ε-nitrogen of K52 on αSyn-V52K and the carboxyl-carbon on the C-terminus of SUMO was below 5 Å, and occasionally coming as close as 3 Å (Figure S20). This suggests that when SUMO was bound to SIM1 of αSyn, its C-terminus can indeed be guided towards V52K of αSyn.

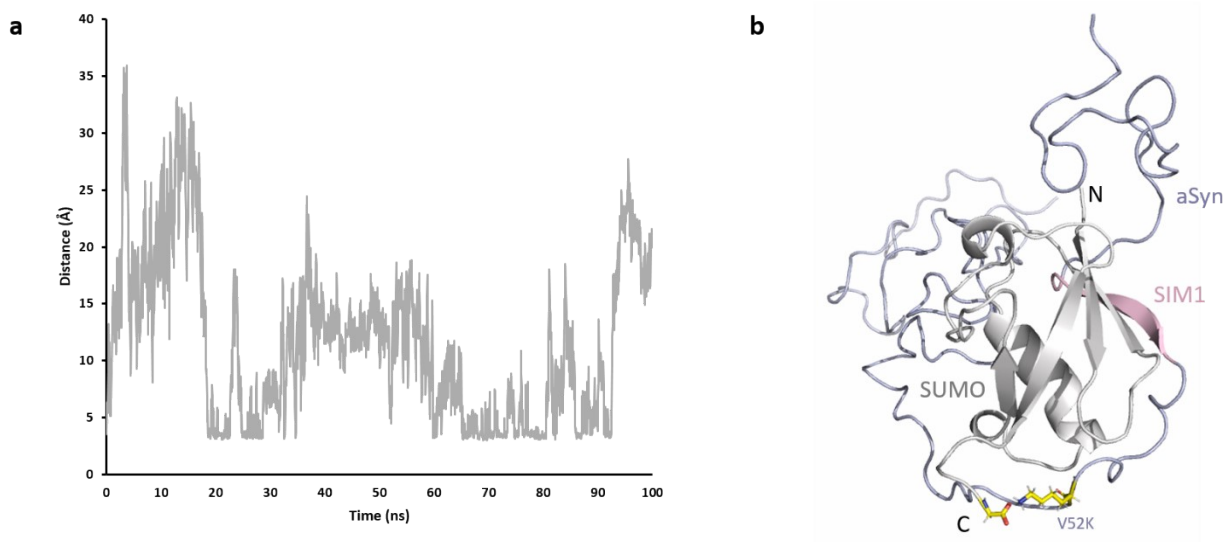


FIGURE S20. Molecular dynamic simulation of αSyn-V52K-SUMO complex. (a) Distance between ε-nitrogen of K52 on αSyn-V52K and the carboxyl-carbon on the C-terminus of SUMO throughout the 100 ns MD simulation. (b) Overall view of the αSyn-V52K-SUMO complex in the simulation snapshot recorded at t=69.1 ns. The SUMO C-terminus was shown to be close to V52K of αSyn when SUMO was bound to SIM1 of αSyn.

To further investigate how the SUMO-SIM interaction contributes to the unexpected enzymatic SUMOylation at K52 of αSyn-V52K (Figure S18), we performed MD simulation of the non-covalent complex between αSyn-V52K, SUMO, and the E2 conjugating enzyme Ubc9 (details in the experimental methods). Examination of the trajectory snapshot revealed that when SUMO was bound to SIM1 of αSyn, both V52K and the SUMO C-terminus could simultaneously fit within the active site of the E2 conjugating enzyme Ubc9 (Figure S21). This suggests that the interaction between SUMO charged on Ubc9 and SIM1 of αSyn can effectively recruit K52 of αSyn-V52K into the active site of Ubc9, thereby facilitating its enzymatic SUMOylation.

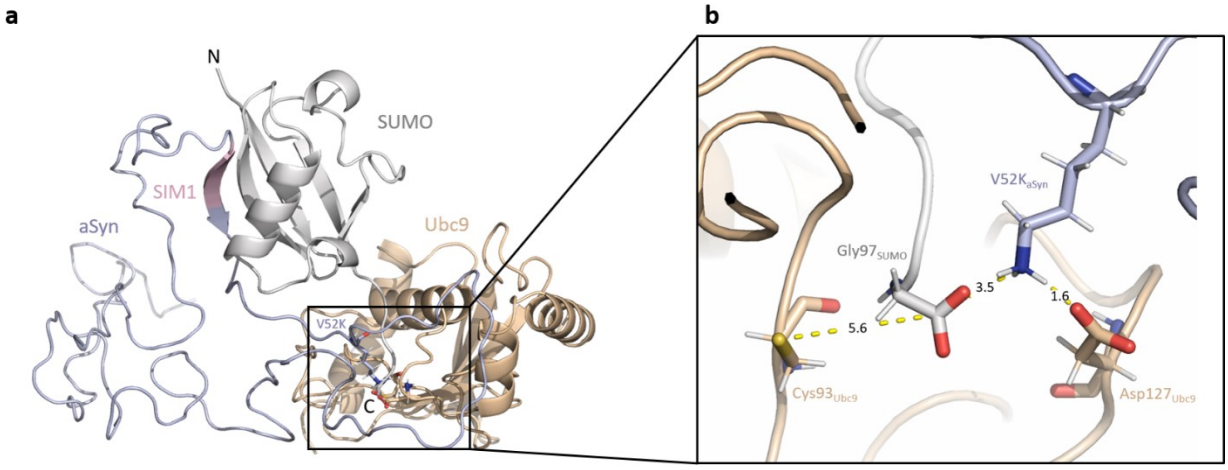


FIGURE S21. Molecular dynamic simulation of αSyn-V52K-SUMO-Ubc9 complex. (a) Overall view of the αSyn-V52K-SUMO-Ubc9 complex in the simulation snapshot at t=42.6 ns. It shows the fitting of V52K and the SUMO C-terminus within the active site of the E2 conjugating enzyme Ubc9 when SUMO was bound to SIM1 of αSyn. (b) Close-up view of the active site of Ubc9 showing the proximity of V52K (αSyn), $Gly9_7$ (SUMO), and the catalytic Cys93 (Ubc9). Note that Asp127(Ubc9) has been proposed to help coordinate and activate the substrate lysine residue.[6]

Supplementary Table

Table S1. Prediction of SUMOylation sites on αSyn-V52K by JASSA.[7] V52K was found to have low predictive score.

Position	Sequence	Results for putatifs SUMO site						
		Best	Consensus direct			Consensus Inverted		
K		PS	Type	PSd	DB Hit	Type	PSi	DB Hit
K6	-----MDVFMKGLSKAKEGVV	None	None	None	1	None	None	
K10	-MDVFMKGLSKAKEGVVAAAE	None	None	None		None	None	
K12	DVFMKGLSKAKEGVVAAAEKT	None	None	None		None	None	1
K21	AKEGVVAAAEKTKQGVAEAAG	None	None	None		None	None	
K23	EGVVAAAEKTKQGVAEAAGKT	None	None	None		None	None	1
K32	TKQGVAEAAGKTKEGVLYVGS	None	None	None		None	None	
K34	QGVAEAAGKTKEGVLYVGSKT	None	None	None		None	None	
K43	TKEGVLYVGSKTKEGVVHGKA	None	None	None		None	None	
K45	EGVLYVGSKTKEGVVHGKATV	None	None	None		None	None	
V52K	SKTKEGVVHGKATVAEKTKEQ	None	None	None		None	None	
K58	VVHGKATVAEKTKEQVTNVGG	None	None	None		None	None	
K60	HGKATVAEKTKEQVTNVGGAV	None	None	None		None	None	
K80	VVTGVTAVAQKTVEGAGSIAA	None	None	None		None	None	
K96	GSIAAATGFVKKDQLGKNEEG	High	NDSM	High	1	None	None	1
K97	SIAAATGFVKKDQLGKNEEGA	None	None	None		None	None	
K102	TGFVKKDQLGKNEEGAPQEGI	None	Weak Consensus	None	1	None	None	1

Amino acid sequences of proteins

(X' denotes the UAG stop codon; The specific amino acid incorporated at this position depends on the non-canonical amino acid supplemented during protein expression)

MmFAcKRS1

MDKKPLNTLISATGLWMSRTGTIHKIKHHEVSRSKIYIEMACGDHLVVNNSRSSRTARALRHHKYRKTCK
RCRVSDEDLNKFLTKANEDQTSVKVKVVSAPTRTKKAMPKSVARAPKPLENTEAAQAQPSGSKFSPAIPV
STQESVSVPASVSTSISSISTGATASALVKGNTNPITSMSAPVQASAPALTKSQTDRLEVLLNPKDEISL
NSGKPFRELESELLSRRKKDLQQIYAEERENYLGKLEREITRFFVDRGFLEIKSPILIPLEYIERMGIDN
DTELSKQIFRVDKNFCLRPMMAPNILNYARKLDRALPDPIKIFEIGPCYRKESDGKEHLEEFTMLNFFQM
GSGCTRENLESIITDFLNHLGIDFKIVGDSCMVYGDTLDVMHGDLELSSAVVGPIPLDREWGIDKPWIGA
GFGLERLLKVKHDFKNIKRAARSESYYNGISTNL*

mCherry-K55UAG-His_6

MVSKGEEDNMAIIKEFMRFKVHMEGSVNGHEFEIEGEGEGRPYEGTQTAKLKVTX'GGPLPFAWDILSPQ
FMYGSKAYVKHPADIPDYLKLSFPEGFKWERVMNFEDGGVVTVTQDSSLQDGEFIYKVKLRGTNFPSDGP
VMQKKTMGWEASSERMYPEDGALKGEIKQRLKLKDGGHYDAEVKTTYKAKKPVQLPGAYNVNIKLDITSH
NEDYTIVEQYERAEGRHSTGGMDELYKLEHHHHHH*

EcADC-His_6

MSKMLRTYNIAWWGNNYYDVNELGHISVCPDPDVPEARVDLAQLVKTREAQGQRLPALFCFPQILQHRLR
SINAAFKRARESYGYNGDYFLVYPIKVNQHRRVIESLIHSGEPLGLEAGSKAELMAVLAHAGMTRSVIVC
NGYKDREYIRLALIGEKMGHKVYLVIEKMSEIAIVLDEAERLNVVPRLGVRARLASQGSGKWQSSGGEKS
KFGLAATQVLQLVETLREAGRLDSLQLLHFHLGSQMANIRDIATGVRESARFYVELHKLGVNIQCFDVGG
GLGVDYEGTRSQSDCSVNYGLNEYANNIIWAIGDACEENGLPHPTVITESGRAVTAHHTVLVSNIIGVER
NEYTVPTAPAEDAPRALQSMWETWQEMHEPGTRRSLREWLHDSQMDLHDIHIGYSSGIFSLQERAWAEQL
YLSMCHEVQKQLDPQNRAHRPIIDELQERMADKMYVNFSLFQSMPDAWGIDQLFPVLPLEGLDQVPERRA
VLLDITCDSDGAIDHYIDGDGIATTMPMPEYDPENPPMLGFFMVGAYQEILGNMHNLFGDTEAVDVFVFP
DGSVEVELSDEGDTVADMLQYVQLDPKTLLTQFRDQVKKTDLDAELQQQFLEEFEAGLYGYTYLEDEHHH
HHH*

EcADC-K240UAG/D535C-His_6

MSKMLRTYNIAWWGNNYYDVNELGHISVCPDPDVPEARVDLAQLVKTREAQGQRLPALFCFPQILQHRLR
SINAAFKRARESYGYNGDYFLVYPIKVNQHRRVIESLIHSGEPLGLEAGSKAELMAVLAHAGMTRSVIVC
NGYKDREYIRLALIGEKMGHKVYLVIEKMSEIAIVLDEAERLNVVPRLGVRARLASQGSGKWQSSGGEX'
SKFGLAATQVLQLVETLREAGRLDSLQLLHFHLGSQMANIRDIATGVRESARFYVELHKLGVNIQCFDVG
GGLGVDYEGTRSQSDCSVNYGLNEYANNIIWAIGDACEENGLPHPTVITESGRAVTAHHTVLVSNIIGVE
RNEYTVPTAPAEDAPRALQSMWETWQEMHEPGTRRSLREWLHDSQMDLHDIHIGYSSGIFSLQERAWAEQ
LYLSMCHEVQKQLDPQNRAHRPIIDELQERMADKMYVNFSLFQSMPDAWGIDQLFPVLPLEGLDQVPERR
AVLLDITCDSDGAICHYIDGDGIATTMPMPEYDPENPPMLGFFMVGAYQEILGNMHNLFGDTEAVDVFVF
PDGSVEVELSDEGDTVADMLQYVQLDPKTLLTQFRDQVKKTDLDAELQQQFLEEFEAGLYGYTYLEDEHH
HHHH*

SUMO(QTC)

MSDQEAKPSTEDLGDKKEGEYIKLKVIGQDSSEIHFKVKMTTHLKKLKESYCQRQGVPMNSLRFLFEGQR
IADNHTPKELGMEEEDVIEVYQEQTC*

ΔN-SUMO(QTC)

MDKKEGEYIKLKVIGQDSSEIHFKVKMTTHLKKLKESYCQRQGVPMNSLRFLFEGQRIADNHTPKELGME
EEDVIEVYQEQTC*

SUMO(KGG)

MSDQEAKPSTEDLGDKKEGEYIKLKVIGQDSSEIHFKVKMTTHLKKLKESYCQRQGVPMNSLRFLFEGQR
IADNHTPKELGMEEEDVIEVYQEQKGG*

TDG(112-360)-His_6

MNGVSEAELLTKTLPDILTFNLDIVIIGINPGLMAAYKGHHYPGPGNHFWKCLFMSGLSEVQLNHMDDHT
LPGKYGIGFTNMVERTTPGSKDLSSKEFREGGRILVQKLQKYQPRIAVFNGKCIYEIFSKEVFGVKVKNL
EFGLQPHKIPDTETLCYVMPSSSARCAQFPRAQDKVHYYIKLKDLRDQLKGIERNMDVQEVQYTFDLQLA
QEDAKKMAVKEEKYDPGYEAAYGGAYGENPCSSEPCGFSSHHHHHH*

TDG(112-360)-K330UAG-His_6

MNGVSEAELLTKTLPDILTFNLDIVIIGINPGLMAAYKGHHYPGPGNHFWKCLFMSGLSEVQLNHMDDHT
LPGKYGIGFTNMVERTTPGSKDLSSKEFREGGRILVQKLQKYQPRIAVFNGKCIYEIFSKEVFGVKVKNL
EFGLQPHKIPDTETLCYVMPSSSARCAQFPRAQDKVHYYIKLKDLRDQLKGIERNMDVQEVQYTFDLQLA
QEDAKKMAVX′EEKYDPGYEAAYGGAYGENPCSSEPCGFSSHHHHHH*

TDG(112-360)-K184UAG-His_6

MNGVSEAELLTKTLPDILTFNLDIVIIGINPGLMAAYKGHHYPGPGNHFWKCLFMSGLSEVQLNHMDDHT
LPGX′YGIGFTNMVERTTPGSKDLSSKEFREGGRILVQKLQKYQPRIAVFNGKCIYEIFSKEVFGVKVKN
LEFGLQPHKIPDTETLCYVMPSSSARCAQFPRAQDKVHYYIKLKDLRDQLKGIERNMDVQEVQYTFDLQL
AQEDAKKMAVKEEKYDPGYEAAYGGAYGENPCSSEPCGFSSHHHHHH*

TDG(112-360)-K206UAG-His_6

MNGVSEAELLTKTLPDILTFNLDIVIIGINPGLMAAYKGHHYPGPGNHFWKCLFMSGLSEVQLNHMDDHT
LPGKYGIGFTNMVERTTPGSKDLSSX′EFREGGRILVQKLQKYQPRIAVFNGKCIYEIFSKEVFGVKVKN
LEFGLQPHKIPDTETLCYVMPSSSARCAQFPRAQDKVHYYIKLKDLRDQLKGIERNMDVQEVQYTFDLQL
AQEDAKKMAVKEEKYDPGYEAAYGGAYGENPCSSEPCGFSSHHHHHH*

TDG(112-360)-K333UAG-His_6

MNGVSEAELLTKTLPDILTFNLDIVIIGINPGLMAAYKGHHYPGPGNHFWKCLFMSGLSEVQLNHMDDHT
LPGKYGIGFTNMVERTTPGSKDLSSKEFREGGRILVQKLQKYQPRIAVFNGKCIYEIFSKEVFGVKVKNL
EFGLQPHKIPDTETLCYVMPSSSARCAQFPRAQDKVHYYIKLKDLRDQLKGIERNMDVQEVQYTFDLQLA
QEDAKKMAVKEEX′YDPGYEAAYGGAYGENPCSSEPCGFSSHHHHHH*

TDG(112-360)-E310Q/K330UAG-His_6

MNGVSEAELLTKTLPDILTFNLDIVIIGINPGLMAAYKGHHYPGPGNHFWKCLFMSGLSEVQLNHMDDHT
LPGKYGIGFTNMVERTTPGSKDLSSKEFREGGRILVQKLQKYQPRIAVFNGKCIYEIFSKEVFGVKVKNL
EFGLQPHKIPDTETLCYVMPSSSARCAQFPRAQDKVHYYIKLKDLRDQLKGIERNMDVQQVQYTFDLQLA
QEDAKKMAVX'EEKYDPGYEAAYGGAYGENPCSSEPCGFSSHHHHHH*

αSyn-His_6

MDVFMKGLSKAKEGVVAAAEKTKQGVAEAAGKTKEGVLYVGSKTKEGVVHGVATVAEKTKEQVTNVGGAV
VTGVTAVAQKTVEGAGSIAAATGFVKKDQLGKNEEGAPQEGILEDMPVDPDNEAYEMPSEEGYQDYEPEA
HHHHHH*

αSyn-K96UAG-His_6

MDVFMKGLSKAKEGVVAAAEKTKQGVAEAAGKTKEGVLYVGSKTKEGVVHGVATVAEKTKEQVTNVGGAV
VTGVTAVAQKTVEGAGSIAAATGFVX'KDQLGKNEEGAPQEGILEDMPVDPDNEAYEMPSEEGYQDYEPE
AHHHHHH*

αSyn-K102UAG-His_6

MDVFMKGLSKAKEGVVAAAEKTKQGVAEAAGKTKEGVLYVGSKTKEGVVHGVATVAEKTKEQVTNVGGAV
VTGVTAVAQKTVEGAGSIAAATGFVKKDQLGX'NEEGAPQEGILEDMPVDPDNEAYEMPSEEGYQDYEPE
AHHHHHH*

αSyn-K10UAG-His_6

MDVFMKGLSX'AKEGVVAAAEKTKQGVAEAAGKTKEGVLYVGSKTKEGVVHGVATVAEKTKEQVTNVGGA
VVTGVTAVAQKTVEGAGSIAAATGFVKKDQLGKNEEGAPQEGILEDMPVDPDNEAYEMPSEEGYQDYEPE
AHHHHHH*

αSyn-V52UAG-His_6

MDVFMKGLSKAKEGVVAAAEKTKQGVAEAAGKTKEGVLYVGSKTKEGVVHGX'ATVAEKTKEQVTNVGGA
VVTGVTAVAQKTVEGAGSIAAATGFVKKDQLGKNEEGAPQEGILEDMPVDPDNEAYEMPSEEGYQDYEPE
AHHHHHH*

αSyn-A91UAG-His_6

MDVFMKGLSKAKEGVVAAAEKTKQGVAEAAGKTKEGVLYVGSKTKEGVVHGVATVAEKTKEQVTNVGGAV
VTGVTAVAQKTVEGAGSIAAX'TGFVKKDQLGKNEEGAPQEGILEDMPVDPDNEAYEMPSEEGYQDYEPE
AHHHHHH*

αSyn-G93UAG-His_6

MDVFMKGLSKAKEGVVAAAEKTKQGVAEAAGKTKEGVLYVGSKTKEGVVHGVATVAEKTKEQVTNVGGAV
VTGVTAVAQKTVEGAGSIAAATX'FVKKDQLGKNEEGAPQEGILEDMPVDPDNEAYEMPSEEGYQDYEPE
AHHHHHH*

αSyn-V95UAG-His_6

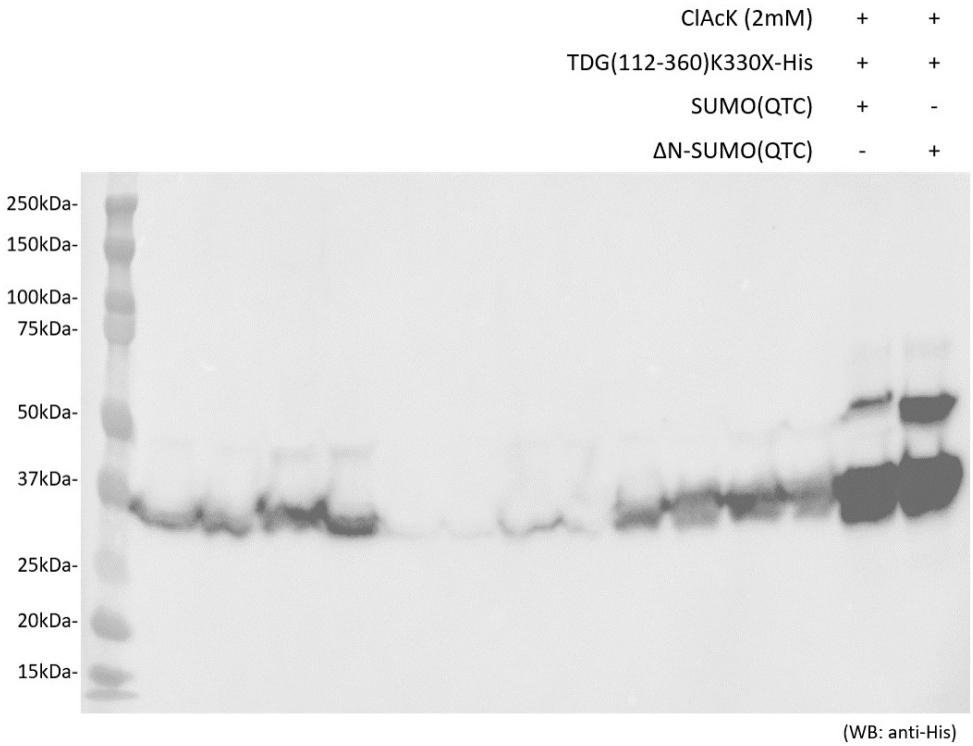
MDVFMKGLSKAKEGVVAAAEKTKQGVAEAAGKTKEGVLYVGSKTKEGVVHGVATVAEKTKEQVTNVGGAV
VTGVTAVAQKTVEGAGSIAAATGFX'KKDQLGKNEEGAPQEGILEDMPVDPDNEAYEMPSEEGYQDYEPE
AHHHHHH*

αSyn-V52K-K96R-K102R-His_6

MDVFMKGLSKAKEGVVAAAEKTKQGVAEAAGKTKEGVLYVGSKTKEGVVHGKATVAEKTKEQVTNVGGAV
VTGVTAVAQKTVEGAGSIAAATGFVRKDQLGRNEEGAPQEGILEDMPVDPDNEAYEMPSEEGYQDYEPEA
HHHHHH*

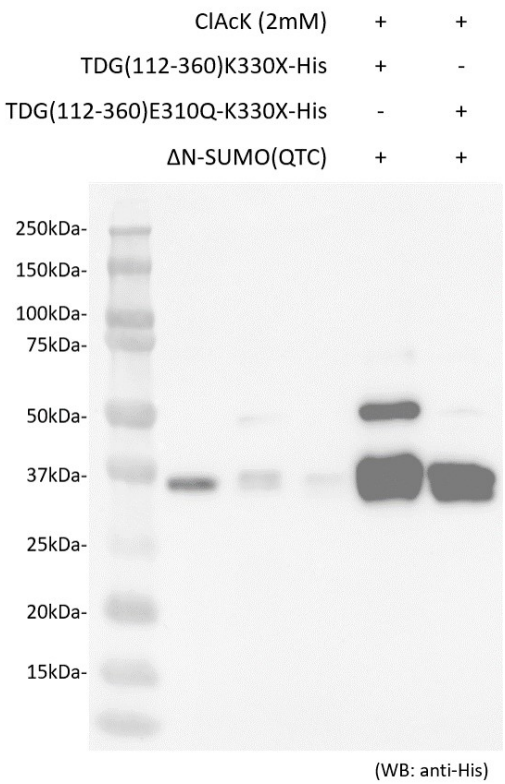
Full gels and blots

Fig. 4a



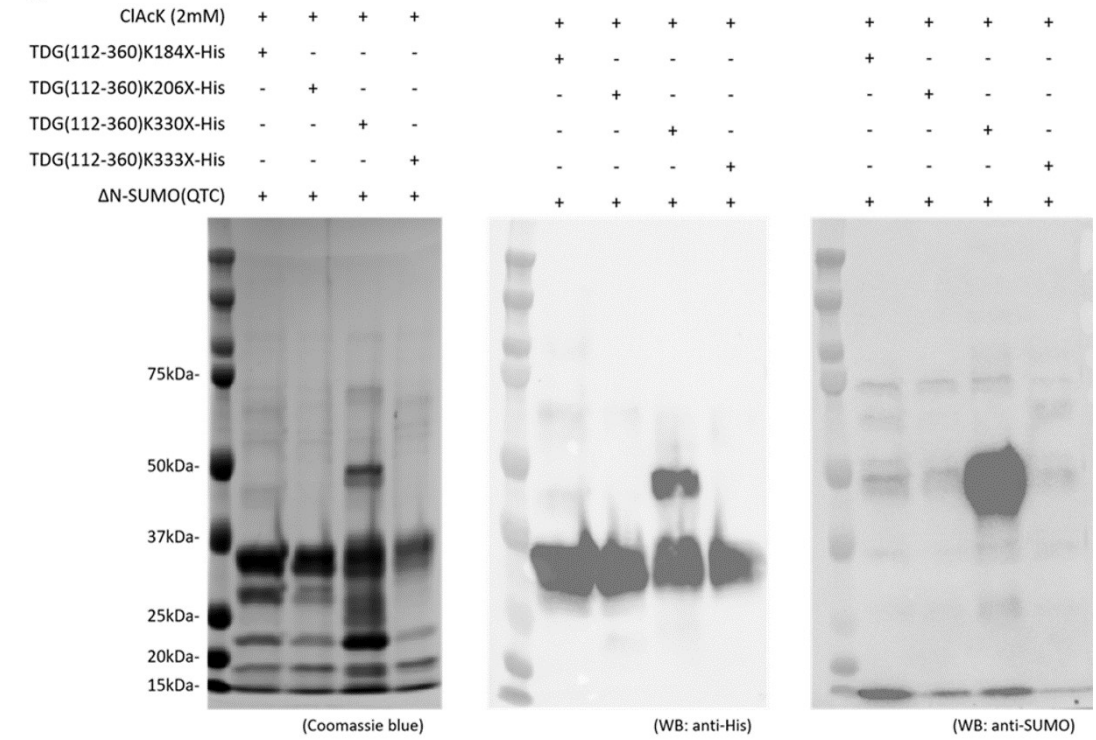
Full blot for Figure 4a.

Fig. 4b



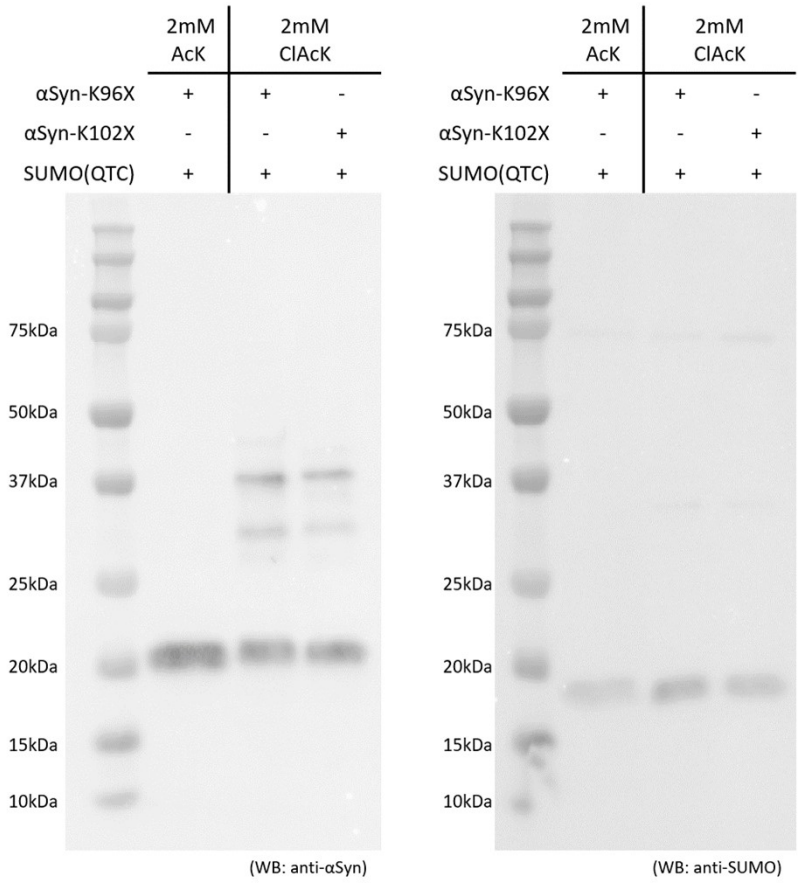
Full blot for Figure 4b.

Fig. 4c

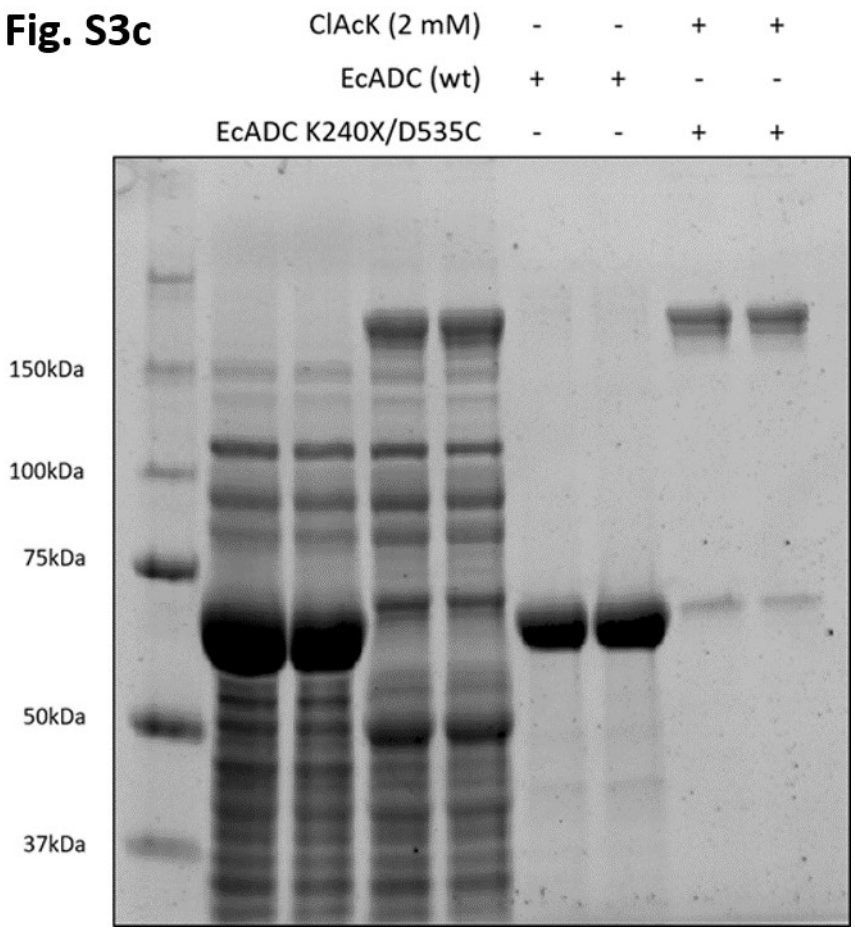


Full blots for Figure 4c.

Fig. 5

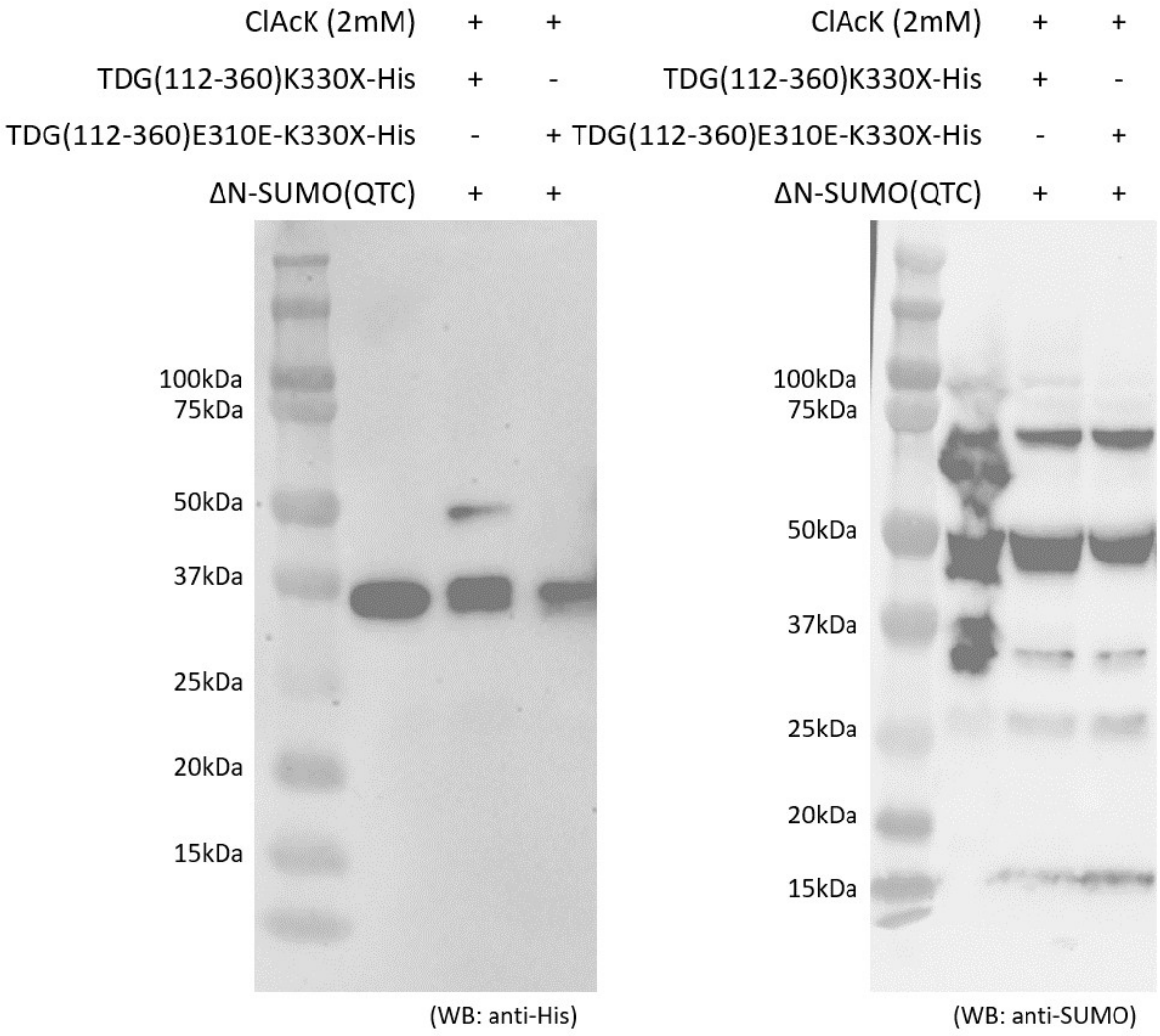


Full blots for Figure 5



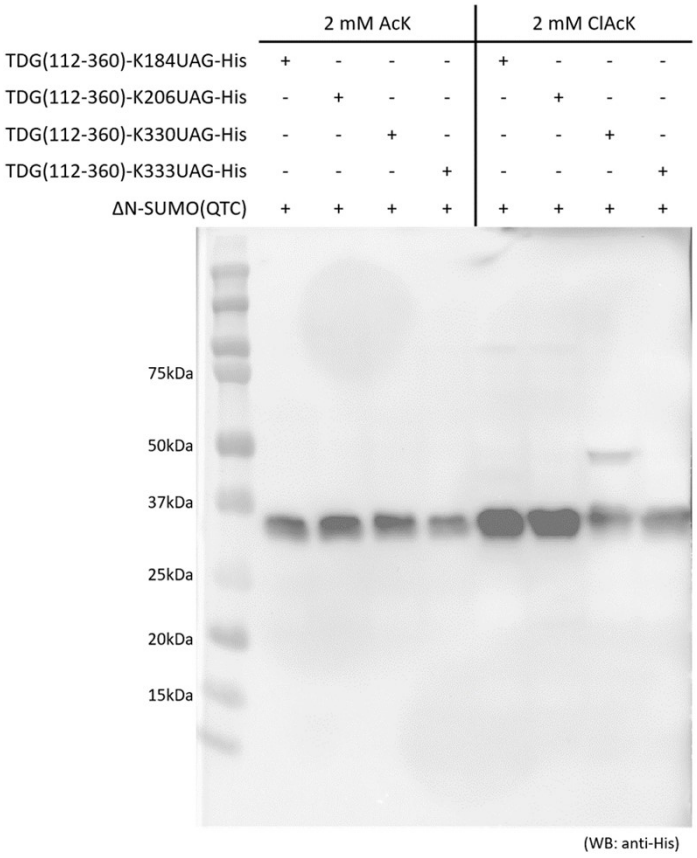
Full gel for Figure S3c

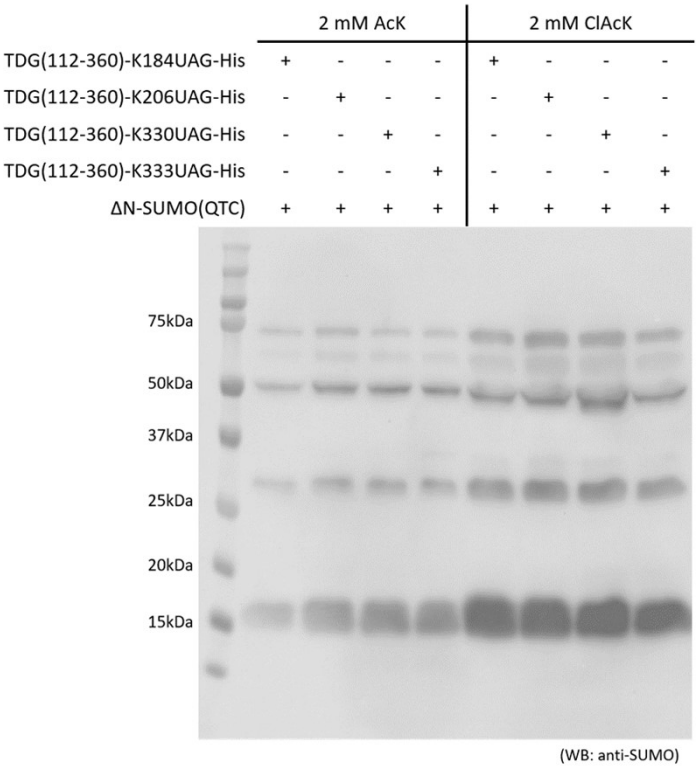
Fig. S6



Full blots for Figure S6

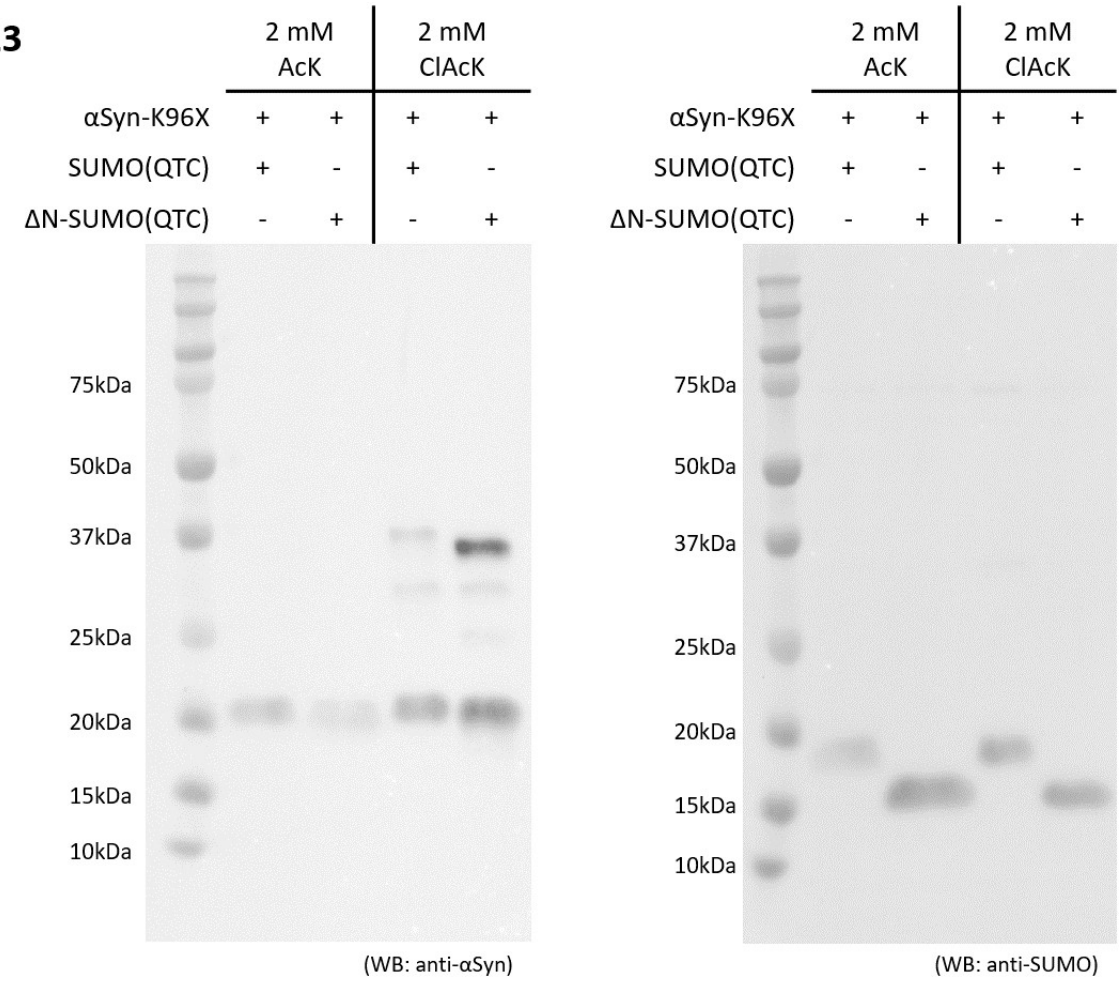
Fig. S7





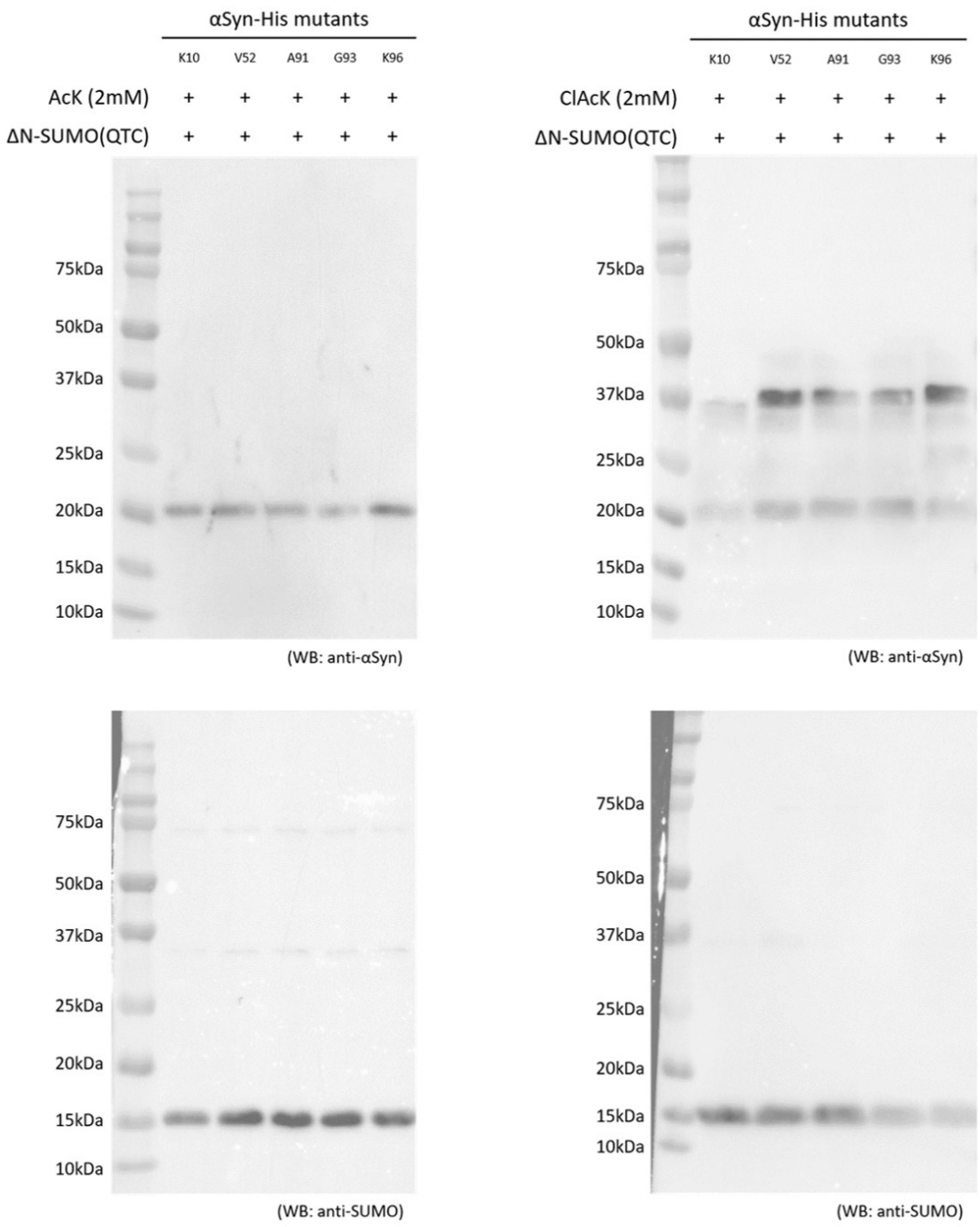
Full blots for Figure S7

Fig. S13



Full blots for Figure S13

Fig. S15



Full blots for Figure S15

References

1. Coey, C. T.; Drohat, A. C., Defining the impact of sumoylation on substrate binding and catalysis by thymine DNA glycosylase. *Nucleic Acids Res.* **2018,** *46* (10), 5159-5170.
2. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A., Highly accurate protein structure prediction with AlphaFold. *Nature* **2021,** *596* (7873), 583-589.
3. Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E., GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015,** *1*, 19-25.
4. Liang, Z.; Chan, H. Y. E.; Lee, M. M.; Chan, M. K., A SUMO1-derived peptide targeting SUMO-interacting motif inhibits α-Synuclein aggregation. *Cell Chem. Biol.* **2021,** *28* (2), 180-190. e6.
5. Meier, F.; Abeywardana, T.; Dhall, A.; Marotta, N. P.; Varkey, J.; Langen, R.; Chatterjee, C.; Pratt, M. R., Semisynthetic, site-specific ubiquitin modification of α-synuclein reveals differential effects on aggregation. *J. Am. Chem. Soc.* **2012,** *134* (12), 5468-5471.
6. Yunus, A. A.; Lima, C. D., Lysine activation and functional analysis of E2-mediated conjugation in the SUMO pathway. *Nature Structural & Molecular Biology* **2006,** *13* (6), 491-499.
7. Beauclair, G.; Bridier-Nahmias, A.; Zagury, J.-F.; Saïb, A.; Zamborlini, A., JASSA: a comprehensive tool for prediction of SUMOylation sites and SIMs. *Bioinformatics* **2015,** *31* (21), 3483-3491.