The proteasomal degradation signal FAT10 modulates substrate stability to accelerate proteasomal degradation

Hitendra Negi[1,2], Aravind Ravichandran[1,2], Pritha Dasgupta[1], Shridivya Reddy[1#], and Ranabir Das[1*]

[1]National Center for Biological Sciences, Tata Institute of Fundamental Research, Bangalore, India

[2]SASTRA University, Thirumalaisamudram, Thanjavur, India

*Electronic address: rana@ncbs.res.in

[#]Currently at Institute for Stem Cell Science and Regenerative Medicine (inStem), Bangalore, India

ABSTRACT

The proteasome machinery regulates the levels of most cellular proteins. However, the factors determining rate of proteasomal degradation remain poorly understood. A rate-limiting step of degradation is the mechanical unfolding of substrates by proteasome ATPases before they enter the core peptidase chamber. Posttranslational modification with the protein FAT10 serves as a proteasomal degradation signal. FAT10 degradation is distinct from the Ubiquitin-proteasome degradation pathway since FAT10 is degraded with the substrate, whereas Ubiquitin is recycled. Hence, both FAT10 and FAT10lyated substrates unfold at the proteasome, and their unfolding kinetics are critical to the degradation rate. Here, we have characterized the thermodynamics and kinetics of unfolding FAT10 and FAT10 conjugated substrate proteins. Our results highlight that fundamental electrostatic interactions of the Ubiquitin fold are absent in FAT10, resulting in a conformationally flexible and malleable native state with low free energy of unfolding. FAT10 substantially modulates substrate stability and induces partially unfolded forms in the substrate to accelerate its unfolding and degradation. The substrate's energy landscape and the FAT10 binding site determine the change in free energy of the substrate. These results highlight a strong dependence of proteasomal degradation rate on the plasticity of proteasomal signal FAT10.

KEYWORDS

Protein degradation, Ubiquitin, FAT10, Proteasome, Proteostasis, NMR spectroscopy

INTRODUCTION

The proteasome is responsible for regulating the levels of a majority of intracellular proteins[1]. Proteins posttranslationally modified with Ubiquitin (Ub) chains interact with the proteasome through Ub receptors in the 19S subunit, which align its base subunits, consisting of AAA+ ATPases, with the 20S Core Particle (CP). The ATPases unfold the substrate and translocate to CP, where the unfolded substrate is further cleaved by peptidases[2–4]. However, the factors determining the efficiency of proteasomal engagement of substrate, and hence, its degradation rate, is poorly understood. Ub is not the sole signal for proteasomal degradation. Another Ubiquitin-like molecule, the human leukocyte antigen-F adjacent transcript 10 (FAT10), directly targets substrates for proteasomal degradation[5]. FAT10 includes two Ub-like domains, an N-terminal domain (FAT10D1) and a C-terminal domain (FAT10D2) attached with a flexible linker. FAT10D1 and FAT10D2 have 29% and 36% sequence similarity to Ub, respectively (Figure 1A). Like Ub, FAT10 is conjugated to substrate by the isopeptide bonds between its C-terminal -GG motif and substrate lysine sidechains. Recent studies suggest that although the fold of FAT10 domains is similar to Ub, their surface properties and interactions with proteasome receptors are distinct [6].

The Ub and FAT10 signaling pathways are diverse in several aspects. While Ub expression is ubiquitous, FAT10 expression is restricted to mature dendritic cells, macrophages, and B cells[8]. Other cell types and tissues express FAT10 upon induction by proinflammatory cytokines [9–11]. FAT10 has a lower half-life (~1 hr) [5] than Ubiquitin (~11 hrs)[7]. FAT10's shorter half-life prevents a prolonged inflammatory response that may lead to apoptosis [12]. Conjugation with a single moiety of FAT10 is sufficient for proteasomal degradation, while tetra-Ub signal is required in the Ub-proteasome pathway. While Ub is cleaved by deubiquitinating enzymes before the substrate enters the 20S core particle, FAT10 is degraded along with the substrate[10]. Typically, the substrate unfolding by the ATPases in the proteasome is a rate-limiting step in substrate turnover[13,14]. However, for the FAT10 substrates, the turnover also depends on the unfolding rate of the proteasome signal FAT10.

Proteasomal substrates need to be globally or partially unfolded for effective engagement with the proteasome. Well folded substrates are preprocessed and extended by cellular unfoldases such as Cdc48/p97/VCP before engaging with the proteasome [2,15]. Recent studies have indicated that a proteasomal tag like Ub can preprocess substrates by destabilizing and partially unfolding them [16–18]. Intriguingly, cellular unfoldases are dispensable for proteasomal degradation of FAT10 substrates [6], alluding to a more substantial effect on substrate energetics by FAT10 compared to Ub. Studying the FAT10 energetics and its impact on the substrate is critical to building a model of the FAT10-proteasome degradation pathway.

Here, we relate the thermodynamic stability and degradation rate of FAT10 and FAT10 substrates. Our results suggest that the energy required to unfold FAT10 is substantially lower than Ub, the kinetics of FAT10 unfolding is seven-fold faster than Ub, and it is easier to unfold FAT10 mechanically. A few salt bridges in Ub increase packing and stabilize its structure. The absence of these salt bridges destabilizes FAT10 and enhances the sampling of partially unfolded forms, leading to faster degradation. FAT10 conjugated substrates are degraded faster than Ub-conjugated substrates. FAT10 significantly reduces the stability of substrate proteins both in cellular conditions and *in-vitro*. FAT10 conjugation modulates the stability of well-folded substrates like Cyan Fluorescent Protein (CFP) by 3 kcal/mol, almost ten-fold higher than Ub. Entropic pulling by FAT10 disrupts native contacts and increases disorder in the substrate, leading to an enhanced sampling of partially unfolded forms. Using NMR spectroscopy, we uncovered that FAT10 modulates energetics in the regions with higher entropy in the free substrate. Additional substrate regions may be perturbed depending on the site of modification. The substrate also reduces the stability of FAT10 by thermodynamic coupling. These results provide the mechanistic understanding underlying the rapid degradation of FAT10 substrates and highlight an intriguing function of FAT10 – to modulate substrate energetics and prime them for degradation.

Materials and Methods

Unfolding studies

Chemical denaturation studies by CD spectroscopy were carried out using 20 µM of protein (Ub, SUMO1, FAT10, FAT10D1, and FAT10D2) incubated overnight with various concentrations of Guanidium Chloride (GdnCl, 0 to 6M) made in the native buffer at 25 ºC. The change in the far-UV signal at 222nm was monitored using the Jasco J-1500 spectropolarimeter. CFP and its variants were studied similarly by incubating 15 nM of protein with GdnCl. CFP fluorescence was measured using Horiba Fluromax-4 with an excitation wavelength of 434 nm. The emission signal collected at 474 nm was averaged for three scans, each for 30 sec, with an integration time of 0.1 sec. Ub_{F45W} and its variants were similarly incubated with GdnCl. The signal at 340 nm was monitored for the change in intrinsic tryptophan fluorescence. The chemical denaturation data were normalized to a two-state unfolding equation using SIGMA plot software. The final curve fitting was done using a monomeric two-parameter melt equation[19]. The calculated free energy of unfolding, $\Delta G_{unfolding}$, and slope of the transition curve, m, were used to calculate free energy (ΔG_0) at 0M [GdnCl] using Eq. 1:

$$\Delta G_{unfolding,\,[GdnCl]} = \Delta G_0 - m[GdnCl] \qquad (1)$$

For thermal equilibrium unfolding studies, the proteins were dialyzed to phosphate buffer (50 mM, pH 6.5, 250 mM NaCl) with or without DTT (1 mM). This unfolding transition was monitored by observing the change in the CD signal at 222nm on the JASCO J-1500 spectropolarimeter connected to a Peltier. The data were collected from 20ºC to 95ºC after every 1ºC per minute rise in temperature with 32 sec of data integration. The raw data of the CD signal were normalized to a two-state unfolding equation and plotted against temperature. The curve was fitted with a sigmoidal equation to yield T_m. Thermal equilibrium unfolding data were analyzed and processed using SIGMA plot software.

Native-state Proteolysis

The protocol for native-state proteolysis has been described previously[16]. 100 µl aliquots of Ub-CFP and FAT10-CFP were treated with multiple thermolysin protease concentrations (stock

concentration 10 mg ml^{-1}). At different time points, 10 µl of the reaction mixture was taken, quenched with 6x-SDS loading dye, and run on SDS-PAGE gels. An I-Bright imaging instrument captured the in-gel fluorescence image (Life Technologies). Images were inverted to greyscale and quantified by normalizing them to no protease control sample. The quantified data were fitted to the first-order exponential equation, and observed proteolysis kinetics (k_{obs}) were calculated at different thermolysin concentrations. The mean of k_{obs} (n=3) at different concentrations were fitted to a linear equation against thermolysin concentration. $\Delta G_{proteolysis}$ was calculated using Eq. 3 by using the slope of the linear fit (Eq. 2). The k_{cat}/K_M of Thermolysin is 99,000M^{-1} s^{-1} [20].

$$\text{Slope of } k_{obs} \text{ vs [Thermolysin]} = K_{op}\,(k_{cat}/K_M) \qquad (2)$$

$$\Delta G_{proteolysis} = -RT \times \ln(K_{op}(k_{cat}/K_M)/99{,}000) \qquad (3)$$

NMR studies

^{13}C,^{15}N isotopically labeled FAT10D1, and FAT10D2 were prepared in 50 mM Tris, 250 mM NaCl buffer with pH 7.5. Standard triple resonance multidimensional NMR experiments ^{15}N-HSQC, HNCA, HN(CO)CA, CBCACONH, and HNCACB were recorded at 298K on an 800MHz Bruker Avance III HD spectrometer equipped with a cryoprobe head. These experiments yielded the backbone assignment of FAT10D1 and FAT10D2. For proteins FAT10D1-Ub and Ub-FAT10D2, only the ubiquitin residue peaks were assigned using standard triple resonance experiments HNCA, HNCOCA, CBCACONH, and HNCACB. NMR data were processed in NMRpipe [21] and analyzed by NMRFAM-SPARKY [22] software. After peak-picking of the backbone experimental data in SPARKY, the peaks were assigned by the PINE server [23] and confirmed manually. The chemical shift perturbations in Ub were calculated between free Ub and FAT10-Ub conjugate as $CSP = \{(\delta N/5)^2 + (\delta H)^2\}^{1/2}$, where δN and δH represent shifts in Nitrogen and Proton dimensions, respectively. The CSPs were calculated similarly between free Ub and FAT10D1-Ub, free Ub and FAT10D2-Ub, and free Ub and Ub-FAT10D2.

Cell Culture Experiments

Approximately 0.1 million cells/well of a 12-well plate were seeded with HEK293T cell line and transfected after 70% confluency using Lipofectamine reagent (Promega). For the cycloheximide assay, cells were transfected with 3xFLAG-wtFAT10-AV, 3xFLAG-wtFAT10D1-CYCIAV, 3xFLAG-wtFAT10D2-AV, and 3xFLAG-UbK0GV. The transfected cells were incubated at 37˚C for 18-20 hours, treated with Cycloheximide (CHX, final concentration 50 µg/mL), and lysed at different time points. For six hours, these cells were treated with MG132 (final concentration 10 µM). 20 mM Tris (pH 7.5), 150 mM NaCl, 1X protease inhibitor cocktail, and 1% NP40 were used to lyse the cell pellet. The total protein amount was estimated by the BCA kit (Thermo) at 37ºC. Similar transfections were carried out for FLAG-UbGV-GFP and FLAG-FAT10AV-GFP. After 18hrs of incubation, the cells were treated with CHX (100 µg/mL) in the presence and absence of p97 inhibitor (DBeQ; final concentration 10µM) and further incubated for different time points. All the Immunoblots were probed with Mouse anti-FLAG antibody (SIGMA, 1 in 10,000 dilutions) and Mouse anti-β-actin antibody (Santa Cruz, 1 in 5000 dilutions) and further probed with HRP conjugated anti-mouse secondary antibody at 1:10,000 dilutions. The blots were developed using clarity ECL (Bio-Rad) staining reagent, and the chemiluminescence signal was observed in Image Quant LAS 4000 (GE). The Immunoblots for GFP variants were probed similarly with the mouse anti-FLAG antibody as above. Tubulin was used as the loading control for normalization and probed with Mouse anti-Tubulin antibody (SIGMA, 1 in 3000 dilutions). All the immunoblots of GFP variants were developed, and the chemiluminescence signal was observed in Amersham Imager 400 (GE). Quantification was done using ImageJ software with three biological replicates for each experiment.

In-cell protein stability assay

The stability of CRABP1 with Ub and FAT10 in the cellular environment was monitored by pre-labeling the *E. coli* BL21 (DE3) cells with FIAsH-EDT_2 fluorescent dye (final concentration 150µM) at OD_{600} = 0.5, followed by incubation till OD_{600} reaches 1. IPTG was added to cultures to induce protein synthesis. After two hours, the culture solution was aliquoted (150 µl) and treated with various

concentrations of Urea ranging from 0M to 3M, keeping the final volume 400 µl. The culture solutions were incubated for 30 mins before pelleting and washing with native buffer (10 mM Tris buffer, pH 7.5), followed by the fluorescent measurements (excitation 500 nm and emission 531 nm). The rest of the protocol is the same as described previously[24].

High-temperature unfolding simulations

Systems were prepared in a cubic box of TIP3p water, with a minimum distance of at least 50 Å between solute atoms and the box edge. Counter ions were added to neutralize the system. The system setup and equilibration procedures were similar to the unrestrained simulations with 450 K as equilibration temperature. Protein unfolding simulations were performed at 450 K for 300 ns. The simulations were performed with ten replicas to obtain an average folding pathway.

ASMD simulations

Constant force explicit solvent ASMD simulations of FAT10D1, FAT10D2, Ub, full-length FAT10, and Ub_2 were carried out. The reaction coordinate is the end-to-end distance between the Cα atoms of the first amino acid and the last amino acid of the respective protein. For simulations consisting of FAT10D1, FAT10D2, and Ub, the proteins were pulled at 1 Å/ns velocity, and for full-length FAT10 and Ub_2, the pulling velocity was 5 Å/ns. The reaction coordinate was partitioned into ten equal segments, each with ten trajectories. The system was energy minimized, equilibrated, and the resulting coordinates & velocities were used as starting points for ASMD simulations. The simulations were performed with periodic boundary conditions in the NPT ensemble; electrostatic interactions were computed by the PME (Particle Mesh Ewald) method. Nonbonded interactions were treated with a cutoff of 8 Å.

Further details of molecular biology, protein purification, MD simulations, and biophysical experiments are provided in the Supporting Information.

Results

FAT10 undergoes rapid proteasomal degradation

The sequence identity between the N- and C-terminal FAT10 domains is poor (18%, Figure 1A). The N-terminal and C-terminal domains will be referred to as FAT10D1 and FAT10D2, respectively (Figure 1B, 1C). The identity of FAT10D1 and FAT10D2 with Ub is 29% and 36%, respectively. The hydrophobic patch L8-I44-V70 is a hot spot of interactions with several Ub co-factors, which is not conserved in the FAT10 domains[6]. A phylogenetic tree analysis based on available ULM structures showed that among the two FAT10 domains, FAT10D2 is structurally closer to Ub (Figure S1A-S1D). Given the poor identity between Ub and FAT10, their cellular and thermodynamic stability could be vastly distinct.

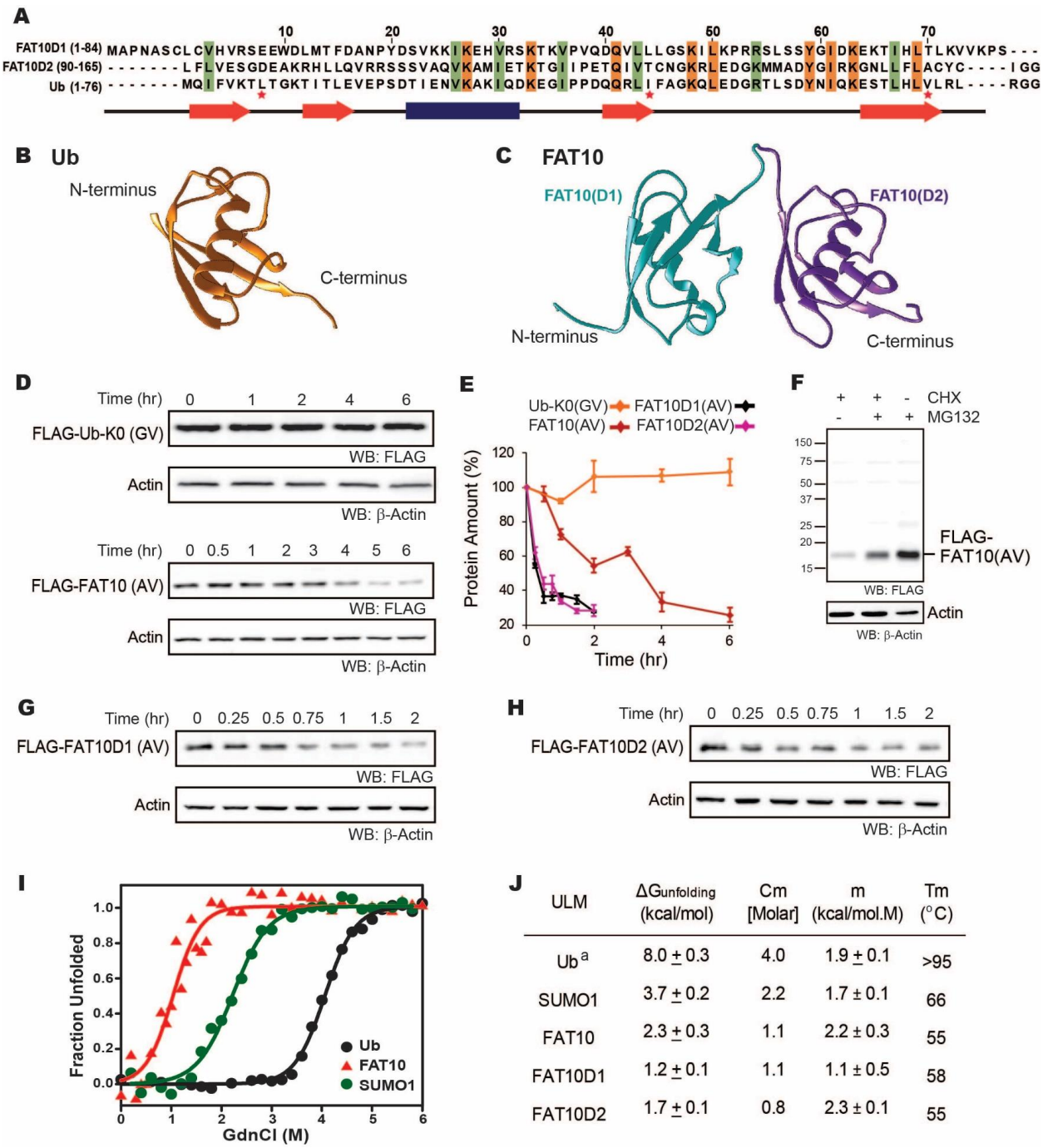
The stability of Ub and FAT10 was compared in cells using cycloheximide chase assay. The non-conjugable UbK0-GV and FAT10-AV were expressed in HEK293T3 cells, and the protein levels were measured against time (Figure 1D). FAT10 levels dropped sharply compared to Ub, suggesting that FAT10 undergoes rapid degradation (Figure 1E). Treatment with proteasomal inhibitor stabilizes FAT10 (Figure 1F), indicating that it undergoes proteasomal degradation. The protein levels of individual FAT10 domains were also studied. Both domains degraded faster than the full-length FAT10 (Figure 1E, 1G, 1H), suggesting that the FAT10 is more stable than the individual domains (Figure 1E). The isolated FAT10 domains were also degraded via the proteasomal pathway (Figure S1E-S1F). The differential stability between the entire FAT10 and the individual domains could be due to the additional unfolding energy required at the proteasome to unfold two domains instead of one. Alternately, transient interactions between the isolated domains may stabilize full-length FAT10.

Thermodynamic characterization of FAT10

It may be argued that the thermodynamics of unfolding Ub and FAT10 is similar, but Ub-specific co-factors stabilize Ub to prevent its degradation in cells. Hence, the energy of unfolding FAT10 and Ub

was measured in a purified system. The far-UV Circular Dichroism (CD) spectra and two-dimensional ^{15}N Heteronuclear Single Quantum Coherence (^{15}N-HSQC) NMR spectra of FAT10 suggested a well-folded tertiary structure (Figure S2A and S2B). The energy of unfolding FAT10 and Ub was measured by chemical denaturation using different quantities of guanidinium hydrochloride (GdnHCl) and monitoring the change in far-UV CD spectra. The Gibbs free energy of unfolding ($\Delta G_{unfolding}$) for Ub was 8 kcal/mol (Figure 1I, 1J). The $\Delta G_{unfolding}$ for another ULM, SUMO1, was lower than Ub (3.7 kcal/mol). However, the $\Delta G_{unfolding}$ was lowest for FAT10 (2.3 kcal/mol, Figure 1I, 1J). Ub has high thermodynamic stability, and its melting temperature (Tm) is greater than 95ºC [25], whereas FAT10 unfolded with a Tm of 55ºC (Figure S2C).

If transient interactions between the FAT10 domains stabilize the full-length protein, the free energy of unfolding individual domains will be lower than the full-length protein[26]. The respective FAT10 domains were isolated, and their energetics was measured. The Far-UV CD and ^{15}N-HSQC spectra confirmed that the purified isolated FAT10 domains are folded (Figure S3A-S3C). GdnHCl-induced and temperature-induced denaturation of the FAT10 domains was studied (Figure S3D-S3F). The $\Delta G_{unfolding}$ was 1.2 kcal/mol and 1.7 kcal/mol for FAT10D1 and FAT10D2, respectively (Figure 1J), which was lower than full-length FAT10, suggesting transient interactions between the FAT10 domains. The difference in $\Delta G_{unfolding}$ and Tm indicates that FAT10 has substantially lower thermodynamic stability than Ub.



ULM	$\Delta G_{unfolding}$ (kcal/mol)	Cm [Molar]	m (kcal/mol.M)	Tm (°C)
Ub[a]	8.0 ± 0.3	4.0	1.9 ± 0.1	>95
SUMO1	3.7 ± 0.2	2.2	1.7 ± 0.1	66
FAT10	2.3 ± 0.3	1.1	2.2 ± 0.3	55
FAT10D1	1.2 ± 0.1	1.1	1.1 ± 0.5	58
FAT10D2	1.7 ± 0.1	0.8	2.3 ± 0.1	55

Figure 1. (A) Structure-based sequence alignment of FAT10D1 (PDB: 1GF1), FAT10D2 (PDB: 1GF2) and Ub (PDB: 1UBQ). Conserved hydrophobic and identical residues in FAT10D1, FAT10D2, and Ub are highlighted in light green and orange colors. The L8-I44-V70 residues that create a 'hot-spot' of interactions in Ub are marked with the red asterisk. (B) Structure of Ubiquitin (1UBQ; orange) and (C)

Homology model structure of full-length FAT10 where FAT10D1 is colored as cyan and FAT10D2 is colored as purple. (D) FLAG-UbK0(GV) and FLAG-FAT10(AV) protein levels are plotted against time. The C-terminal GG residues are substituted with GV or AV to prevent conjugation to the cellular substrates. HEK293T cells were transfected with either FLAG-Ub or FLAG-FAT10(AV), treated with Cycloheximide, and lysed at different time points. The lysates were separated on SDS PAGE gels and blotted with anti-FLAG antibodies. (E) Quantified protein levels in (D) are plotted against time after normalizing with β-Actin. (F) HEK293T cells were transfected with FAT10, treated with/without Cycloheximide, and proteasomal inhibitor MG132. The lysates were separated on SDS PAGE gels and blotted with anti-FLAG antibodies. (G) Similar to (D), showing degradation of FLAG-FAT10D1(AV) & (H) FLAG-FAT10D2(AV) after cycloheximide treatment. Quantified protein levels of FLAG-FAT10D1(AV) and FLAG-FAT10D2(AV) are plotted in (E). (I) GdnCl melt curves of FAT10, Ub, and SUMO1. Normalized mean ellipticity shift is plotted against GdnCl concentration. (J) A table with details of the stability parameters of Ub, SUMO1, FAT10, and FAT10 domains is provided. (a: Refer Wintrode et al., 1994).

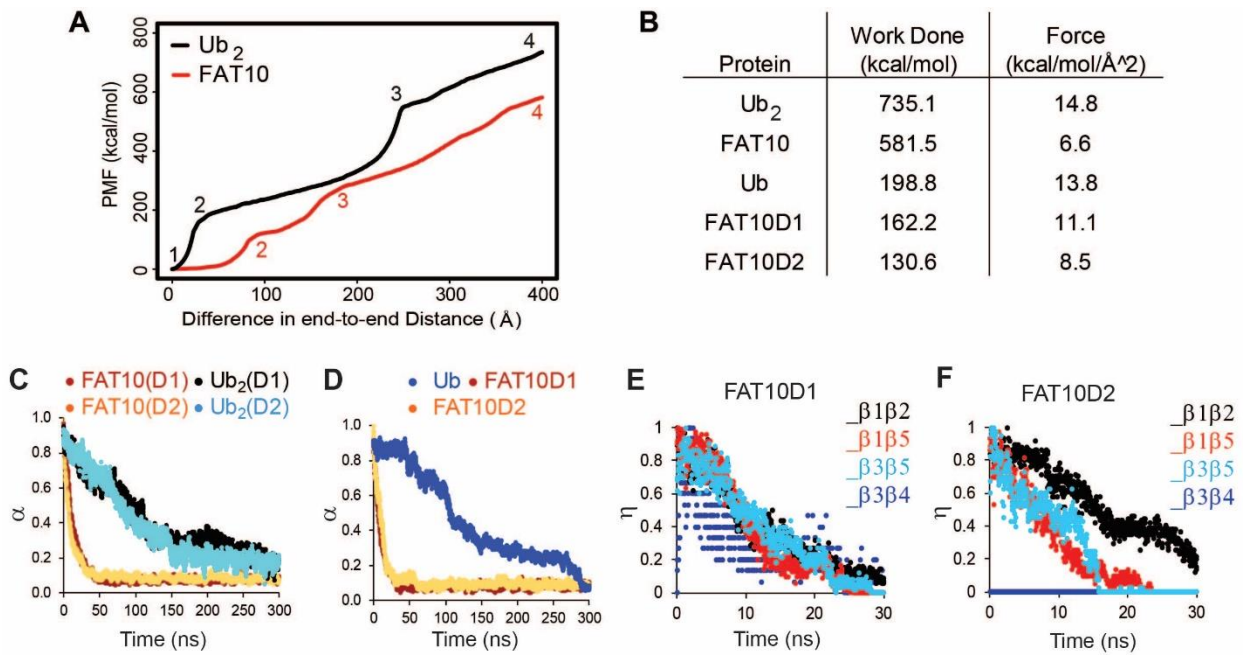
The distinct unfolding pathway of FAT10

The mechanical unfolding of proteins by molecular motors such as the proteasomal ATPases can be simulated by steered molecular dynamics[27]. The mechanical unfolding of FAT10 was studied in the explicit solvent by Adaptive Steered Molecular Dynamics (ASMD) (Figure 2A, Movie S1, Figure S4)[28]. For comparison, a linear Di-Ubiquitin (Ub_2) was used, whose size is similar to FAT10. While the work required to unfold and linearize Ub_2 was 735 kcal/mol, the same for FAT10 was 581 kcal/mol (Figure 2B). The lower work required to unfold FAT10 corroborated the lower value of $\Delta G_{unfolding}$ measured during FAT10 denaturation (Figure 1J). A similar simulation was repeated for each isolated FAT10 domain and compared to mono Ub. The work required to unfold the isolated FAT10 domains was lower than Ub by 30-60 kcal/mol (Figure 2B, Figure S5). Resistance during mechanical unfolding arises from

cooperative packing interactions between buried sidechain atoms[29]. The lower resistance in FAT10 suggests weaker cooperative packing interactions at the buried core of the protein.

The kinetics of FAT10 unfolding was studied by high-temperature all-atom molecular dynamics simulations. The simulations were carried out in FAT10 and Ub_2 for a total time of 4 μs (10 x 400 ns) at 450K (Movie S2). The enthalpy and stability of proteins are governed by the strength of Van der Waals (VdW) interactions between nonpolar groups and Hydrogen bond (Hbond) interactions between polar groups. As the molecules unfold, the native intramolecular contacts, including VdW interactions and Hbonds, decrease with time. Whereas the interactions in Ub_2 decreased gradually over 300 ns, they reduced drastically by 40 ns in FAT0, suggesting weaker packing interactions in FAT10 (Figure 2C). The ULM proteins adopt a β-grasp fold consisting of a five-strand β-sheet. The β-sheet Hbonds disrupted promptly in FAT10 as opposed to Ub (Figure S6A). The native contacts and Hbonds in the individual isolated FAT10 domains were also disrupted faster than Ub (Figure 2D and S6B).

The fraction of native Hbonds between each β-strand pair in the β-sheet of FAT10 domains and Ub were plotted (Figure 2E, 2F, and S6C). The sequence in which such pairs disrupt is distinct between the FAT10 domains and Ub, suggesting different unfolding pathways. A comprehensive analysis of the secondary structure Hbonds during unfolding showed that the 3^{10} helix α2 unfolds first in Ub, followed by β3-β4, α1, β3-β5 and β1-β5 contacts (Figure S6D). The β1-β2 contacts break at the last step. The Ub unfolding pathway is similar to that observed in previous equilibrium simulations[30]. In contrast, the first unfolding event in FAT10D1 is the disruption of β3-β4 contacts, followed by α2, α1, and β1-β5 contacts (Figure S6E). The β3-β5 and β1-β2 contacts are disrupted simultaneously at the last step. The first two unfolding steps in FAT10D2 are similar to FAT10D1 (Figure S6F). However, the β3-β5 contacts break prematurely while β1-β2 persist longer. Altogether, the work done to mechanically unfold FAT10 is lower than Ub, the kinetics of FAT10 unfolding is faster than Ub, and the FAT10 unfolding pathway is distinct from Ub.



Protein	Work Done (kcal/mol)	Force (kcal/mol/Å^2)
Ub_2	735.1	14.8
FAT10	581.5	6.6
Ub	198.8	13.8
FAT10D1	162.2	11.1
FAT10D2	130.6	8.5

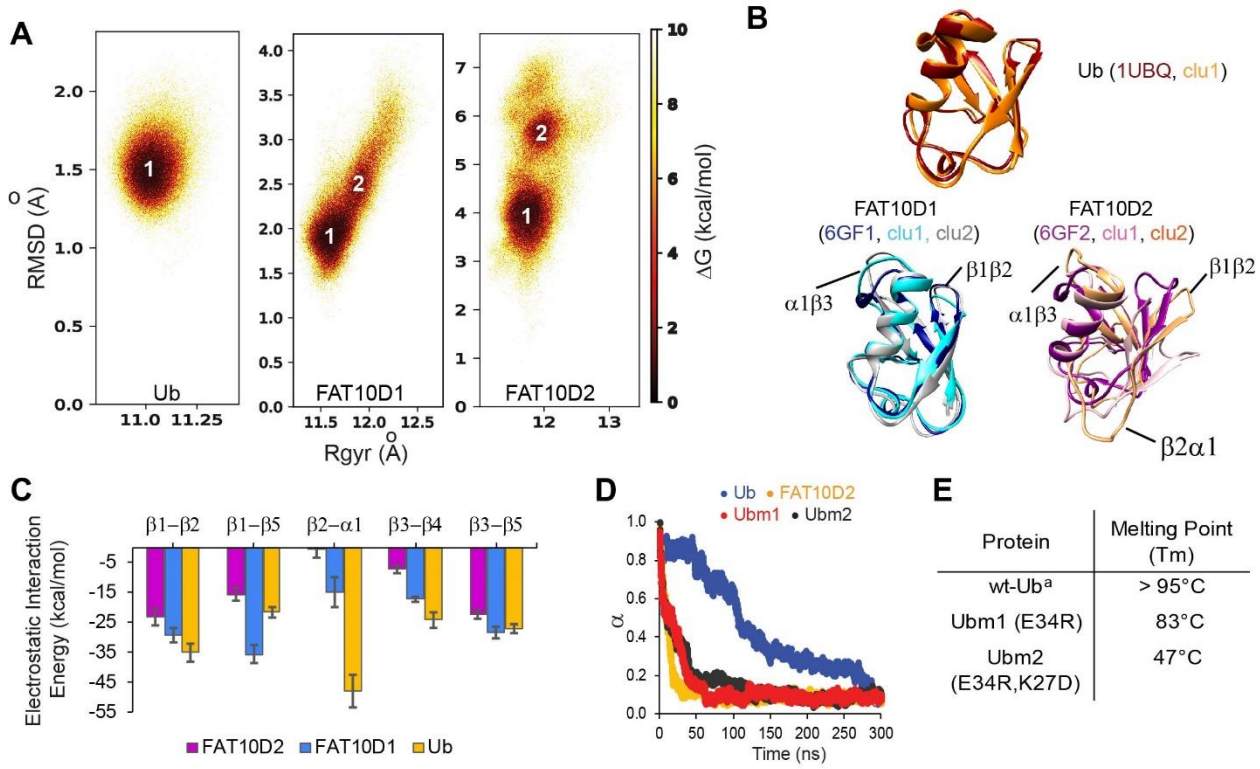
Figure 2. Unfolding studies of FAT10, di-Ubiquitin (Ub_2), Ub, and individual domains D1/D2 in FAT10 by MD simulations. (A) ASMD of Ub_2 and FAT10. The potential Mean Force (PMF) is plotted against the normalized end-to-end distance. Each unfolding event is marked by a number, whose corresponding conformation is given in Figure S4. (B) The work done in unfolding the proteins by ASMD is provided. (C) Simulations of FAT10, di-Ubiquitin (Ub_2), Ub, and individual domains D1/D2 in FAT10 were performed at 450K. The fraction of native contacts is defined as (α) and plotted against time. FAT10(D1) and FAT10(D2) are the D1 and D2 domains in FAT10. Ub_2(D1) and Ub_2(D2) are the two Ub domains in diubiquitin. The data is averaged over ten replicas. (D) is the same as (C) measured for individual domains FAT10D1, FAT10D2, and Ub. (E) The fraction of native pairwise backbone hydrogen bond is defined as (η), which is plotted for β1 to β5 of FAT10D1. (F) The same as (E) for the FAT10D2 domain.

Key electrostatic interactions stabilize Ub but not FAT10

Partially unfolded forms[31] in the substrate and proteasomal tag can facilitate engagement with proteasome[16]. The dynamics of FAT10 domains and Ub were studied by MD simulations at room temperature (300 K) for 7.5 μs (3 x 2.5 μs). The Root Mean Square Fluctuations (RMSF) of Ub are lower than FAT10 domains at the β1-β2 loop, C-terminal end of α1, and the α2 loop (Figure S7A, S7B). The free energy of protein conformations was calculated as a function of root mean square deviation (RMSD) and the radius of gyration (Rgyr). The 2D density plot of the free energy of Ub reflected a single highly populated state (Figure 3A). The representative structure of the state superimposes well with the native state Ub X-ray structure (Figure 3B). In contrast, the ground states of FAT10 domains are in equilibrium with higher energy partially unfolded forms. The loops at N- and C-termini of α1 are partially unfolded compared to the ground state. The interactions between β1-β2 and α1 are disrupted, which resulted in altered structure is these regions.

The radius of gyration (Rog) reflects the nature and extent of packing interactions present in the protein. Although Ub and the FAT10 domains are similar in size, the Rog of Ub ground state is 11.9 Å, 12.4 Å for FAT10D1, and 12.6 Å for FAT10D2 (Figure S7C). The higher Rog values for FAT10 domains suggest that they are less compact than Ub. To understand the basis of reduced compactness and packing interactions, the VdW interaction energies and the electrostatic interaction energies were calculated between the interacting secondary structure elements in FAT10 and Ub. The VdW interaction energy was comparable between FAT10 domains and Ub (Figure S7D). The electrostatic interaction energy was also similar in all regions except between β2 and α1, where Ub has a significantly higher negative value suggesting extensive electrostatic contacts in this region (Figure 3C). Three hydrogen bonds and a salt bridge between β2 and α1 are present in Ub but absent in FAT10, which could be responsible for the lower electrostatic energy (Figure S8A). Another salt bridge between the α2 loop and α1 is present in Ub but absent in FAT10 (Figure S8B). These salt bridges could be crucial to the thermodynamic stability of Ub.

When the salt bridges were disrupted in Ub by substitution, its free energy landscape changed, giving rise to partially unfolded states similar to FAT10 (Figure S9A). The RMSF values of mutant Ub increase at the C-terminal end of α1 and the loop between β4 and α2 loop (Figure S9B, S9C). Principle Component Analysis (PCA) and the RMSF values suggested that the Ub mutants have increased dynamics in the regions of broken salt bridges (Figure S9D-S9F). When the unfolding kinetics were compared between Ub mutants and wtUb by simulations at 450K, the Ub mutants unfolded much faster than Ub but similar to FAT10 domains (Figure 3D). The thermodynamic stability of Ub mutants was measured experimentally. Disrupting the salt bridges reduces the melting point of Ub by 50°C, suggesting the importance of the salt bridges for Ub stability (Figure 3E and Figure S9G). The salt bridges were engineered in the FAT10D1 domain by appropriate substitutions, and their effect on the FAT10D1 was investigated. The engineered salt bridges removed the partially unfolded states in FAT10D1 and created a single state (Figure S10A). RMSF analysis and PCA analysis suggested that engineered salt bridges reduced the fluctuations in FAT10D1 (Figure S10B-S10D). Overall, critical electrostatic interactions present between α1 and β1β2, and between α1 and α2 loop stabilize the Ub fold. These interactions are absent in FAT10, resulting in weaker intramolecular contacts, increasing flexibility and rate of unfolding.



Protein	Melting Point (Tm)
wt-Ub[a]	> 95°C
Ubm1 (E34R)	83°C
Ubm2 (E34R,K27D)	47°C

Figure 3. (A) The free energy landscape of Ub, FAT10D1, and FAT10D2 plotted as a function of RMSD and radius of gyration (Rgyr) obtained from simulations across three replicas (3 x 2.5 μs) performed at 300K. The minima from each cluster are numbered. (B) The corresponding conformation of each cluster in (A) is shown. The structures of the proteins, denoted by their PDB ids, are provided for comparison. (C) The mean electrostatic energy of interactions between different pairs of secondary structures obtained from simulations is plotted for Ub, FAT10D1, and FAT10D2. The error bars denote standard deviation. (D) The fraction of native contacts in Ubm1, Ubm2, Ub, and FAT10D2 against time at 450K simulations. (E) The melting point of Ubiquitin mutants, where the salt bridges are disrupted. (a: Refer Wintrode *et al.*, 1994).

Weaker hydrogen bonds and higher conformational entropy in FAT10

The factors governing the protein stability and dynamics include the enthalpy of interactions and the conformational entropy. The enthalpy of interactions consists of the VdW and Hbond interactions.

While experimentally measuring the strength of VdW interactions is challenging, the temperature dependence of backbone amide proton chemical shifts is an excellent reporter of Hbond strength[32]. The temperature dependence of the amide proton chemical shifts is referred to as the temperature coefficient Tc, where Tc = $\Delta\delta^{NH}/\Delta T$ [33]. Backbone amides form Hbond with other protein atoms or with solvent, and Tc measures the strength of backbone Hbonds at high resolution. Higher negative Tc values suggest weaker Hbonds and disorder propensity, while lower negative Tc values suggest stronger Hbonds and rigidity[32]. Tc's were measured in FAT10 and plotted against each residue in Figure 4A (fitting errors are plotted in Figure S11A). In the D1 domain in FAT10, residues in the region β1 to β2 (β1β2), and the loop between β4 and β5 (α2 loop) had high negative Tc values (Figure 4A). Similarly, several residues in the β1, β2, α1, and α2 loop in the D2 domain had high negative Tc values, suggesting weaker Hbonds in these regions.

Tc's were measured in Ub and plotted against each residue in Figure 4B (fitting errors in Figure S11B). The loop between β1 and β2 had high negative Tc values in Ub, which correlates with the region's higher mobility and disorder propensity[34]. Due to the low sequence similarity between FAT10 domains and Ub, it is challenging to compare the Tc values between individual residues of the two proteins. Instead, Tc values were averaged over the different segments in the proteins and compared (Figure S11G). The difference in averaged Tc values between the D1 domain, FAT10(D1), and Ub indicate that the β1β2 region and the α2 loop have weaker Hbonds in D1 (Figure 4C, 4D). In the D2 domain, the β1β2 regions and α1 had weaker Hbonds than Ub. The Tc measurements were repeated for the individual isolated FAT10 domains (FAT10D1 and FAT10D2) and analyzed similarly (Figure 4E, Figure S11E, S11F, S11H). The β5 Hbonds were more robust in the full-length FAT10 than the individual domains, suggesting that the transient interactions with the other domain may stabilize β5.

To estimate the differences in conformational entropy between FAT10 and Ub, the backbone order parameters (S^2) were measured by standard NMR relaxation experiments. The spin-lattice relaxation (R1), spin-spin relaxation (R2), and heteronuclear NOEs (hetNOE) were measured in FAT10

and Ub (Figure S12). S^2 values were calculated using the Lipari-Szabo Model-Free Analysis method and averaged over the protein segments. The difference in averaged S^2 between FAT10 and Ub was converted to conformational entropy $-T\Delta S$[35,36] and plotted in Figure 4F-H. The higher value of $-T\Delta S$ suggests higher dynamics and flexibility. The β1β2 and α2 loop had high conformational entropy in FAT10(D1) domain compared to Ub (Figure 4F, 4G). In the FAT10(D2) domain, β1β2, β3β4, and helix α1 were more entropic compared to Ub, the most significant being α1 (Figure 4G, 4H). Similar measurements were performed in the isolated individual FAT10D1 domain (Figure S13). The β1β2 and α2 loop had higher mobility than the rest of the protein, similar to the full-length FAT10 (Figure S13C). Overall, Ub salt bridges increase interactions between α1 and β1β2, and between α1 and the α2 loop. In their absence, these regions in FAT10 have weaker Hbonds and higher conformational entropy.

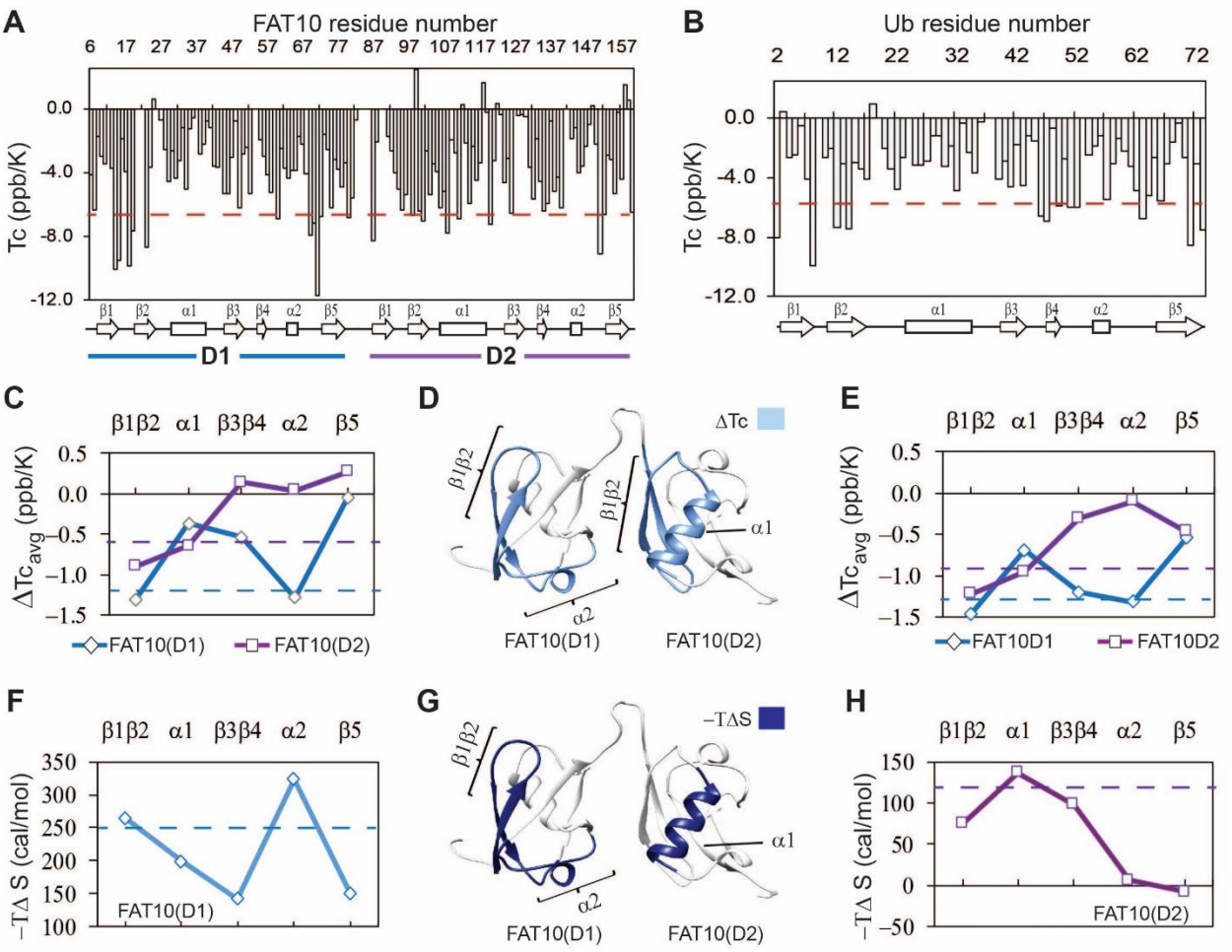
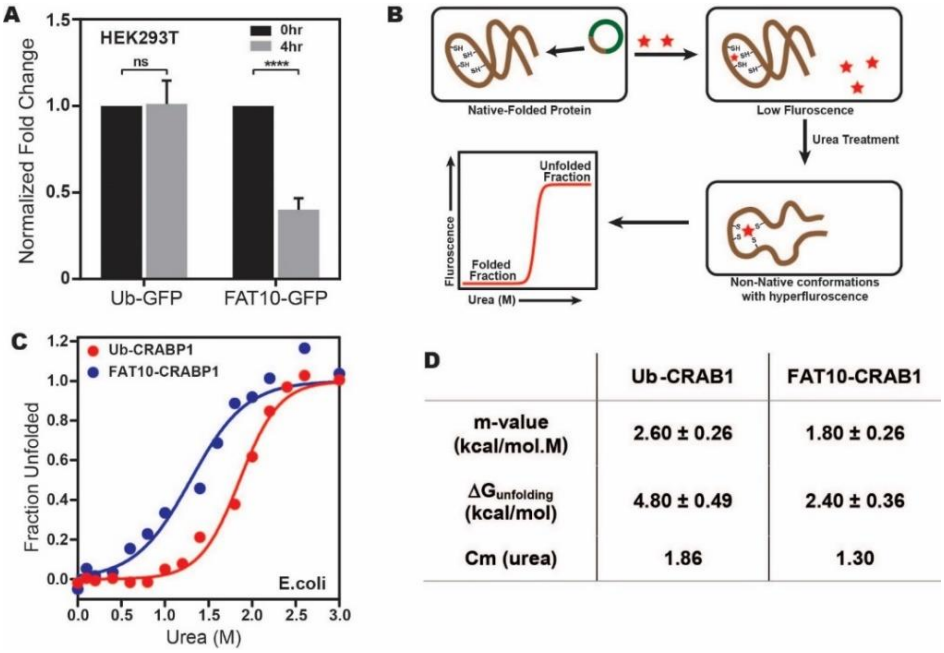


Figure 4. A comparison of the local stability and dynamics between FAT10 and Ub. The temperature dependence of amide chemical shifts is plotted for FAT10 and Ub in (A) and (B), respectively. Temperature coefficients (Tc) are plotted against the residue numbers of each protein. The horizontal red line is (mean – S.D.), where the mean value is negative. High negative Tc values suggest weaker Hbonds and disorder propensity. (C) The difference in averaged temperature coefficients (ΔTc_{avg}) between FAT10 and Ub, where $\Delta Tc_{avg} = Tc(Ub)_{avg} - Tc(FAT10(D1))_{avg}$. The blue line is (mean –SD) for FAT10(D1). The purple line is the (mean –SD) for FAT10(D2). Higher negative ΔTc_{avg} values suggest weaker Hbonds and destabilization in these FAT10 regions compared to Ub. (D) The segments with negative ΔTc values are colored light blue on the FAT10 structure. (E) The same as (C) plotted for isolated individual domains FAT10D1/FAT10D2. The horizontal blue and purple lines are (mean –SD) for FAT10D1 and FAT10D2,

respectively. (F) The difference in conformational entropy $-T\Delta S_{conf}$, where $\Delta S_{conf} = S_{conf}^{Ub} - S_{conf}^{FAT10(D1)}$, was averaged for the various segments and plotted. The entropy values were calculated from the order parameters measured in Figure S12. The broken line denotes (mean + SD). (G) The segments with $-T\Delta S$ more than (mean + SD) are colored blue on FAT10(D1). (H) The difference in conformational entropy between Ub and FAT10(D2) is plotted similarly to (F). The segments with $-T\Delta S$ more than (mean + SD) are colored blue on FAT10(D2) in (G). Higher values of $-T\Delta S$ suggest increased conformational flexibility in these FAT10 regions compared to Ub.

FAT10 destabilizes substrates in cellular conditions

The degradation rate of FAT10 substrates depends on how FAT10 modulates the substrate energetics. The stability of FAT10 conjugated substrates was investigated in cells by a cycloheximide assay using GFP as the model. FAT10 conjugated GFP was degraded to 40% within 4 hours after cycloheximide treatment, much faster than the degradation of Ub conjugated GFP (Figure 5A, S14A). However, it remains unclear if FAT10 actually affects substrate energetics in cellular conditions or merely serves as a better tag with a higher affinity for proteasomes. In bacteria, the Pup-proteasome system functions by a distinct mechanism[37] that does not recognize the eukaryotic Ub or FAT10, making it easier to study changes in substrate stability by FAT10/Ub in cellular conditions. The model substrate CRABP1 was chosen for studying the effect of FAT10 on the substrate stability in *Escherichia coli* cells. CRABP1 can be engineered to bind a fluorescent dye that is quenched in the native state but not in the unfolded state [24] to study the folded to the unfolded transition of CRABP1 (Figure 5B). Cells expressing CRABP1 were treated with various Urea concentrations. The difference in fluorescent signals between 0M and 3M urea was lowest in apo CRABP1, higher in Ub-CRABP1, and highest in FAT10-CRABP1, suggesting the CRABP1 stability is lowest when conjugated to FAT10 (Figure S14B-S14D). The $\Delta G_{unfolding}$ was 2.4 kcal/mol less in FAT10-CRABP1 than Ub-CRABP1 (Figure 5D), which suggested significant substrate destabilization by FAT10 in cellular conditions.



	Ub-CRAB1	FAT10-CRAB1
m-value (kcal/mol.M)	2.60 ± 0.26	1.80 ± 0.26
$\Delta G_{unfolding}$ (kcal/mol)	4.80 ± 0.49	2.40 ± 0.36
Cm (urea)	1.86	1.30

Figure 5. (A) Comparison of Ub-GFP and FAT10-GFP levels post 4hrs of treatment with Cycloheximide (CHX) in HEK293T cells. (B) The schematic for studying the substrate stability in a heterologous cellular system, where a protein (CRABP1) capable of binding FIAsH-EDT2 dye was expressed in *E.coli* BL21 (DE3) cells as a substrate. The folded protein quenches the dye, while unfolded protein releases the quenching. (C) Fluorescence signals of Ub-CRABP1 and FAT10-CRABP1 were normalized to plot their denaturation curves. (D) Thermodynamic parameters of CRABP1 in cellular conditions, when covalently bound to Ub and FAT10, respectively.

FAT10 modulates the substrate energetics

In highly crowded cellular environments, proteins experience various physical interactions, such as the exclusion volume effect and nonspecific transient interactions. To study whether the substrate destabilization is solely due to FAT10 and not the above effects, the impact of FAT10 on the substrate energetics was studied in purified systems. Producing purified FAT10-conjugated substrates is technically

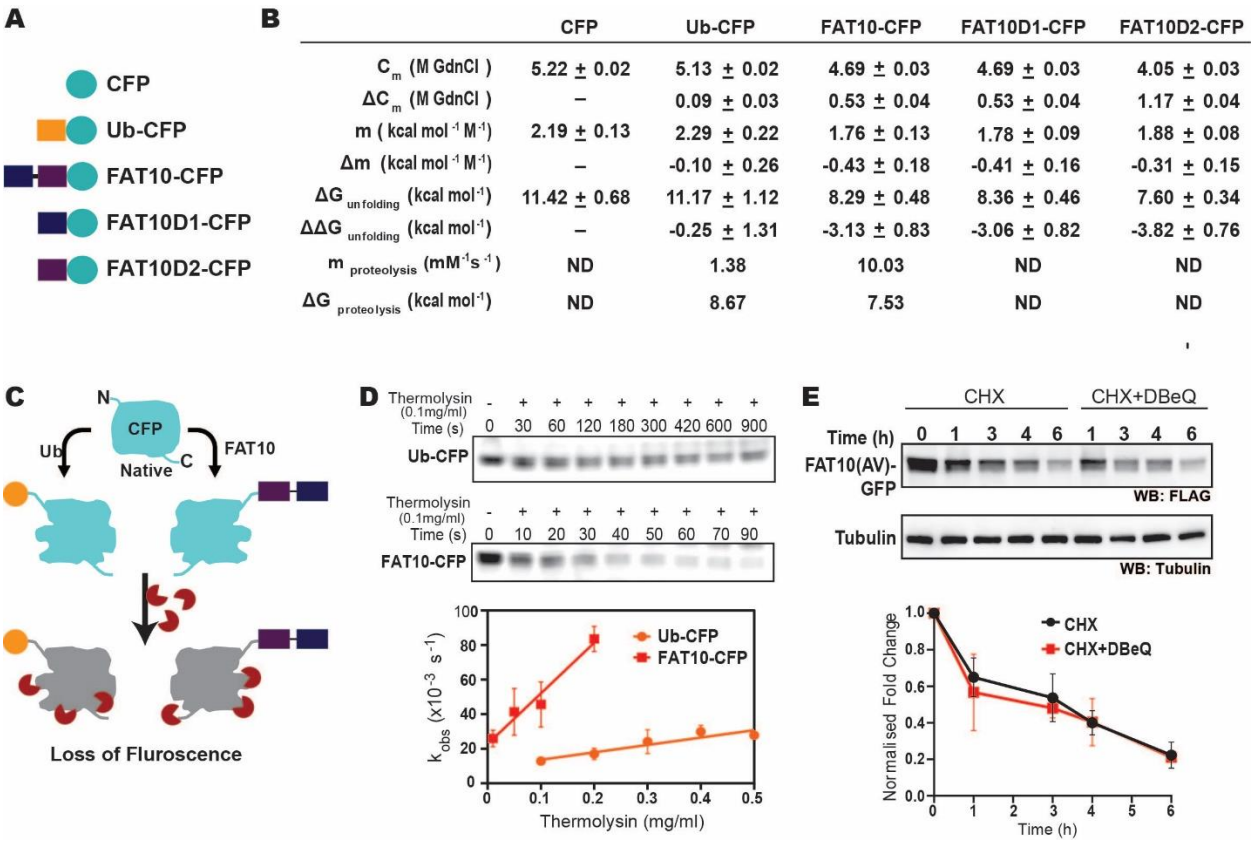
challenging because a FAT10 E3-substrate pair active under *in vitro* conditions are yet to be identified. Hence, FAT10 was covalently conjugated at the N-terminus of a model ultra-stable substrate protein like the Cyan Fluorescent Protein (CFP) to mimic the N-terminal modification[38] (Figure 6A). The free energy of unfolding ($\Delta G_{unfolding}$) of CFP was exclusively monitored by measuring the CFP fluorescence during chemical denaturation. The change in $\Delta G_{unfolding}$ in CFP is modest ($\Delta\Delta G_{unfolding}$ = 0.25 kcal/mol) when CFP is covalently linked with Ub (Figure 6B and Figure S15A). The modest decrease in CFP stability correlates with the finding that Ub needs cellular unfoldases to degrade well-folded substrates[15]. Interestingly, $\Delta\Delta G_{unfolding}$ of CFP was 3.1 kcal/mol when covalently linked to FAT10, suggesting FAT10 destabilizes CFP ten-fold more than Ub (Figure 6B and Figure S15A). The free energy change could be due to the entropic 'pulling' effect[39] or specific interactions between FAT10 and CFP.

Specific interactions should depend on the complementary surface properties of FAT10 and CFP and alter if the FAT10 is truncated to smaller individual domains. The stability of CFP was monitored by N-terminal conjugation with the individual domains FAT10D1 and FAT10D2 (Figure 6A). The $\Delta\Delta G_{unfolding}$ was 3.0 kcal/mol and 3.8 kcal/mol when CFP is conjugated to FAT10D1 and FAT10D2, respectively (Figure 6B and Figure S15B). The $\Delta\Delta G_{unfolding}$ due to the isolated domains is similar to the full-length FAT10, suggesting that instead of specific interactions, entropic pulling by nonspecific interactions may destabilize CFP.

Under cellular conditions, proteins tend to have more partially than globally unfolded conformations. The partially unfolded conformations are critical to forming effective interactions with the proteasome[16]. The reduced stability in FAT10 substrates suggests that FAT10 may enhance sampling partially unfolded forms in the substrate. The susceptibility of partially unfolded conformations to a typical protease like thermolysin was measured in CFP, Ub-CFP, and FAT10-CFP (Figure 6B-6D, Figure S15C-F). Although CFP was stable in the presence of thermolysin, both Ub-CFP and FAT10-CFP were proteolyzed. FAT10-CFP proteolysis rate (k_{obs}) was significantly more than Ub-CFP at a given thermolysin concentration (Figure 6D). The change in observed proteolysis rates with different

thermolysin concentrations ($m_{proteolysis}$) was seven-fold higher in FAT10-CFP than Ub-CFP (Figure 6D). The values of $m_{proteolysis}$ can be correlated to the free energy of proteolysis[20]. The free energy of proteolysis decreased from 8.7 kcal/mol for Ub-CFP to 7.5 kcal/mol for FAT10-CFP (Figure 6B), indicating that FAT10 substantially increases the population of partially unfolded states in CFP.

Our results indicate FAT10 has a more drastic effect than Ub on substrate stability, and FAT10 substrates could be degraded independently of the unfoldases. The degradation of FAT10-GFP was monitored in the HEK293T cells in the presence and absence of the Cdc48 inhibitor DBeQ (Figure 6E). The rate of FAT10-GFP degradation was unperturbed by DBeQ, suggesting FAT10 modified substrates are degraded independently of Cdc48. Overall, FAT10 modulates the substrate energetics significantly higher than Ub.



	CFP	Ub-CFP	FAT10-CFP	FAT10D1-CFP	FAT10D2-CFP
C_m (M GdnCl)	5.22 ± 0.02	5.13 ± 0.02	4.69 ± 0.03	4.69 ± 0.03	4.05 ± 0.03
ΔC_m (M GdnCl)	–	0.09 ± 0.03	0.53 ± 0.04	0.53 ± 0.04	1.17 ± 0.04
m (kcal mol^{-1} M^{-1})	2.19 ± 0.13	2.29 ± 0.22	1.76 ± 0.13	1.78 ± 0.09	1.88 ± 0.08
Δm (kcal mol^{-1} M^{-1})	–	-0.10 ± 0.26	-0.43 ± 0.18	-0.41 ± 0.16	-0.31 ± 0.15
$\Delta G_{unfolding}$ (kcal mol^{-1})	11.42 ± 0.68	11.17 ± 1.12	8.29 ± 0.48	8.36 ± 0.46	7.60 ± 0.34
$\Delta\Delta G_{unfolding}$ (kcal mol^{-1})	–	-0.25 ± 1.31	-3.13 ± 0.83	-3.06 ± 0.82	-3.82 ± 0.76
$m_{proteolysis}$ ($mM^{-1}s^{-1}$)	ND	1.38	10.03	ND	ND
$\Delta G_{proteolysis}$ (kcal mol^{-1})	ND	8.67	7.53	ND	ND

Figure 6. (A) Different versions of CFP used in this study where CFP is fused at the C-terminal end of Ub, FAT10, FAT10D1, and FAT0D2. (B) Table with thermodynamic and proteolysis parameters of CFP bound to FAT10 domains are provided. (C) Native state proteolytic cleavage of Ub-CFP and FAT10-CFP using thermolysin. (D) Representative in-gel fluorescence image for native-state proteolysis of Ub-CFP and FAT10-CFP at 0.1 mg ml^{-1} thermolysin. k_{obs} of proteolysis are plotted for Ub-CFP and FAT10-CFP against various thermolysin concentrations. Error bars denote the standard deviation of the replicates (n=3). (E) *In-vivo* degradation of FAT10-GFP after cycloheximide treatment without/with p97 inhibitor (DBeQ). Tubulin is used as the loading control. The quantified level of FLAG-FAT10-GFP without/with inhibitor is given below.

FAT10 destabilizes areas with higher entropy in the substrate

Whether FAT10 changes the substrate energetics globally or locally at specific regions is unknown. Ub was chosen as a model substrate to study the FAT10 effect at a higher resolution. Covalent interactions between FAT10 and Ub have been reported previously[40]. Simulations studied the impact of FAT10 on the kinetics of Ub unfolding at 450K. The C-terminal Ub in diUb unfolded similarly to monoUb (Figure S16A and S16C). However, the Ub in FAT10D1-Ub or FAT10D2-Ub unfolded faster than monoUb (Figure S16B and S16D), suggesting that FAT10 domains increase the rate of Ub unfolding. The FAT10 domains also enhanced dynamics in the helix α1 and the loop between β4 and α2 in Ub (Figure 7A), and induced partially unfolded forms (Figure 7B).

FAT10's effect on substrate energetics was experimentally measured using $Ub_{45F/W}$ mutant as a substrate, where the intrinsic tryptophan fluorescence reports the thermodynamics of substrate unfolding. $Ub_{45F/W}$ is active, and its thermodynamic stability is similar to Ub[41]. The single tryptophan in the N-terminal domain of FAT10 is mutated to Phenylalanine (F) to not interfere with the fluorescence spectra (Figure S16E). The free energy of unfolding Ub reduced from 11.7 kcal/mol to 7 kcal/mol when

covalently bound at the N-terminus with FAT10 (Figure 7C and Figure S16F-H). The individual domains of FAT10 had a similar effect on Ub. The domain D2 was conjugated to the C-termini of Ub (Ub-FAT10D2) to examine the site-specific impact on the substrate. The $\Delta\Delta G_{unfolding}$ in Ub was higher when the D2 domain was at the C-terminal end than at the N-termini (Figure 7C, Ub-FAT10D2 versus FAT10D2-Ub), suggesting the substrate energetics may be regulated by site-specificity.

The conformational space of the proteasomal tag (FAT10 domains) about the substrate (Ub) and substrate/tag contacts was examined to investigate the site-specific effects further. Several nonspecific contacts were observed between the tag and the substrate, including transient Hbonds and VdW interactions (Figure S17). The conformational space sampled by the tag FAT10D2 in Ub-FAT10D2 was higher than FAT10D2-Ub (Figures 7D and 7E). Consequently, the number of nonspecific associations between the substrate and tag was also higher in Ub-FAT10D2 (Figure S17G), which may lead to lower $\Delta\Delta G_{unfolding}$ and higher destabilization. The substrate can also influence the stability of the proteasomal tag by thermodynamic coupling. The free energy surface of free FAT10D2 is indeed more compact than the FAT10D2 in FAT10D2-Ub and Ub-FAT10D2 (Figure S18A). Moreover, the unfolding of FAT10D2 is faster when covalently bound to Ub (Figure S18B), suggesting a thermodynamic coupling between the FAT10 tag and the substrate.

The changes in the local environment in Ub due to the FAT10 and its domains were measured by NMR. ^{15}N-edited HSQC spectra of the conjugated proteins showed well-dispersed backbone amide resonances, suggesting that they are folded (Figure S19). Chemical shifts of backbone amide resonances are excellent reporters of changes in the local chemical environment. The Chemical shift perturbation (CSP) between free Ub and the FAT10-conjugated Ub shows perturbation in the β1, β2, and β2α1 loop (Figure S20A and S20C). For FAT10D1-Ub, the perturbations were distributed in the β1, β2, α1, and α2 loops of Ub (Figure S20B and S20D). For FAT10D2-Ub, the CSPs were localized to the β1, β2, and α2 loop (Figure 7F, Figure S20E). The CSPs were similar when FAT10D2 was present at the C-terminus, except that the α2 loop CSPs were significantly higher (Figure 7G, Figure S20F). Given that C-terminally

tagged Ub is more unstable, the higher CSPs in the α2 loop suggest that changes in this region can substantially affect Ub stability. Altogether, the chemical environment in the substrate is not perturbed globally but in specific areas when conjugated with FAT10.

To measure FAT10's effect on the strength of Hbonds in the substrate, Tc values of Ub residues in FAT10D2-Ub, Ub-FAT10D2, and FAT10D1-Ub were measured (Figure S20G). Due to a severe overlap in the NMR spectra, Tc values in FAT10-Ub could not be measured. The differences in averaged Tc values between FAT10D2-Ub and free Ub show that the Hbonds in the strands β1β2, helix α1 and the α2 loop are weakened in Ub (Figure 7H). Additional Hbonds are destabilized in β3β4, and β5 when FAT10D2 is conjugated at the C-terminus (Figure 7H), which is commensurate with its lower $\Delta G_{unfolding}$. The effect in β5 is probably due to the covalent binding of FAT10D2 at the C-terminal end of β5. The β4 is closely connected to β5 by multiple Hbonds, and perturbation of β5 also affects β3β4. Changes in the substrate's conformational entropy were quantified by measuring the order parameter of Ub in FAT10D2-Ub and Ub-FAT10D2 (Figure S21). The difference in order parameters between free Ub and Ub in the conjugates were converted into conformational entropy (Figure 7I). N-terminal covalent binding of FAT10D2 increases entropy at the β1β2 loop, β2, and α1 in Ub. However, entropy increases further in these regions when FAT10D2 is covalently bound at the C-terminus. In addition, the α2 loop and β3β4 regions become more entropic. Overall, FAT10 domains weaken Hbonds and increase conformational entropy at specific regions in Ub. The strength of these effects is dependent on the site of covalent binding.

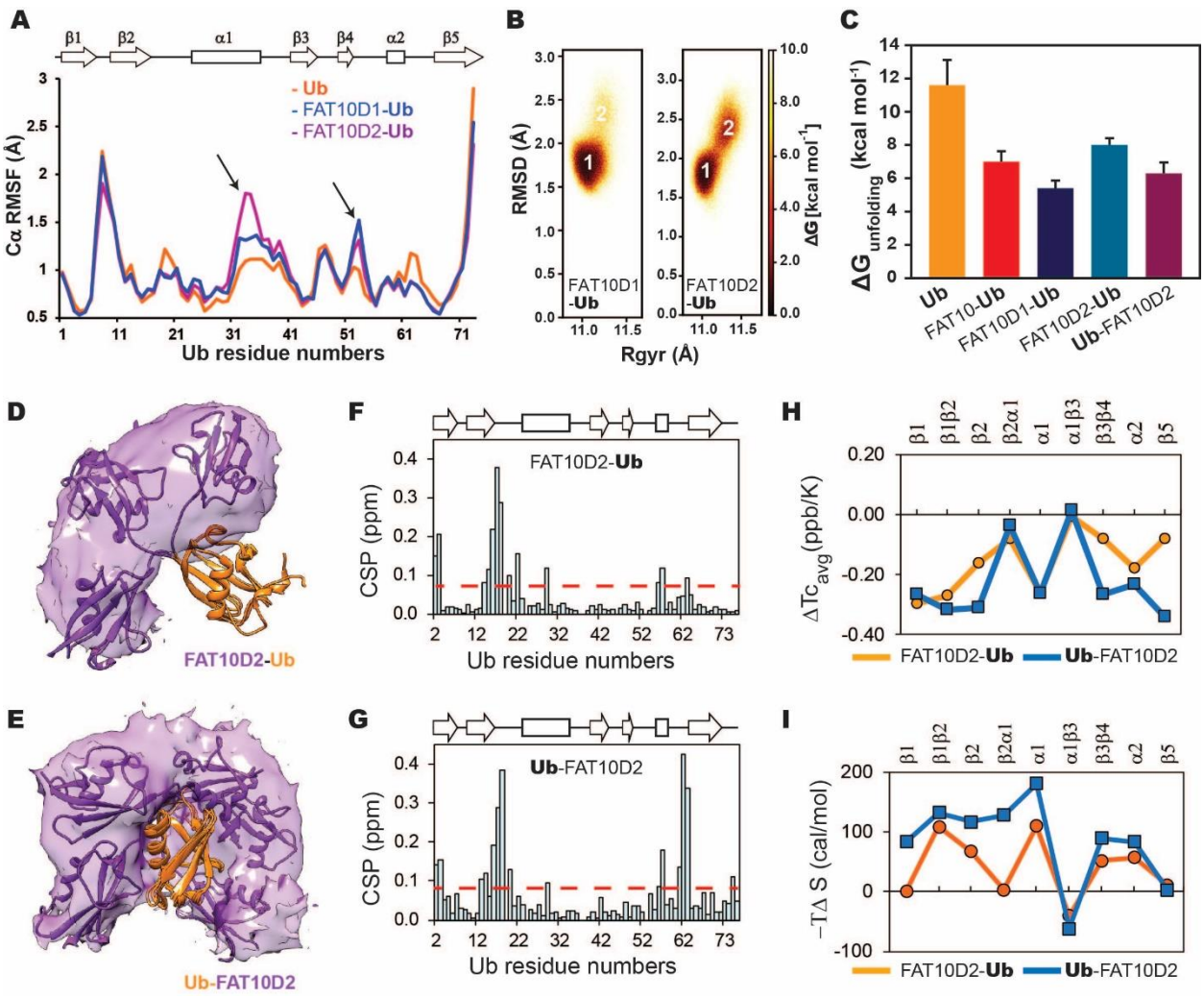


Figure 7. (A) A comparison of Cα RMSF values of Ub in mono-Ub, FAT10D1-Ub, and FAT10D2-Ub. (B) The free energy landscape of Ub in FAT10D1-Ub and FAT10D2-Ub are plotted as a function of RMSD and radius of gyration (Rgyr) obtained from simulations across three replicas (3 x 2.5 μs) performed at 300K. (C) The free energy of unfolding $Ub_{45F/W}$ in free form and covalently bound to FAT10 domains is plotted. The occupancy of the FAT10D2 around Ub is shown as a purple surface for (D) FAT10D2-Ub and (E) Ub-FAT10D2. Few structures during the simulation are shown. The CSPs in Ub due to covalent attachment with the FAT10D2 domain for (F) FAT10D2-Ub and (G) Ub-FAT10D2. The broken red line denotes SD. (H) The difference in mean temperature coefficients between FAT10D2-Ub and free Ub is plotted for the various Ub segments. The orange line is ΔTc_{avg}=Tc(FAT10D2-Ub) $_{avg}$ -

$Tc(Ub)_{avg}$. The blue line is the same for Ub-FAT10D2, where $\Delta Tc_{avg}=Tc(Ub\text{-}FAT10D2)_{avg}-Tc(Ub)_{avg}$. (I) The difference in conformational entropy $-T\Delta S$ of Ub between free Ub and FAT10D2-Ub, where $\Delta S = S^{Ub} - S^{FAT10D2\text{-}Ub}$ was averaged for the various segments and plotted (orange). The same was plotted for Ub-FAT10D2 in blue.

Discussion

The mechanical unfolding of proteins at the proteasome is energy-consuming, rate-limiting, and regulates the substrate turnover [13,14]. Since FAT10 is degraded with the substrate, the substrate turnover depends on both FAT10 and substrate energetics. The free energy of FAT10 unfolding is significantly lower than Ub by 6 kcal/mol, and the kinetics of FAT10 unfolding is tenfold faster than Ub. Unlike Ub, the native state of FAT10 domains exists in equilibrium with partially unfolded forms that accelerate proteasomal degradation. FAT10 weakly resists mechanical unfolding, reflecting weak packing interactions within the molecule. Our results uncover that the lack of strong electrostatic interactions between the central helix α1 and β1β2 strands, and the α1 and α2 loop, creates a malleable structure in FAT10 with lower enthalpy of interactions and higher conformational entropy than Ub. Such a structural design expedites the unfolding at proteasome and accelerates the degradation of FAT10 substrates.

Using a variety of substrates, we observed that FAT10 reduces the stability of its substrates, both *in-vitro* and in cellular conditions. We chose two ultra-stable proteins, CFP and Ub, whose melting points are above 90°C. FAT10 destabilized both the substrates by 3-5 kcal/mol, significantly higher than Ub. Substrate destabilization either by complete FAT10 or its individual domains was comparable, suggesting that destabilization is not dependent on the size but composition of the degradation tag. The exposed hydrophobic surface area is 23% in each FAT10 domain and 13% in Ub. The significantly exposed hydrophobic surface in FAT10 can induce nonspecific interactions to destabilize the substrate (Figure S17). Moreover, the FAT domains tend to sample partially unfolded forms, exposing their buried

hydrophobic residues, further enhancing the nonspecific interactions between the tag and the substrate. The mechanism is reminiscent of prokaryotic and eukaryotic chaperones that unfold substrates by multiple nonspecific interactions for translocation across organelle membranes[42]. Additionally, we noted that the substrate affects the energy landscape of the proteasomal tag, creating partially unfolded forms in the tag. Hence, thermodynamic coupling between the substrate and FAT10 will destabilize both in tandem and promote degradation.

An increase in conformational entropy associated with reduced enthalpy of interactions in folded proteins/domains suggests order-to-disorder transition and reduced stability. Enthalpic changes are primarily due to the disruption of hydrogen bonds, monitored here by measuring the changes in the temperature dependence of backbone amide NMR chemical shifts (ΔTc). The changes in enthalpy and conformational entropy were correlated in the substrate upon FAT10 interaction (Figure 7). The conformational entropy increased in the regions where the enthalpic component decreased, indicating reduced stability. FAT10 induced the sampling of partially unfolded forms in the substrate, as observed in proteolytic cleavage assays and simulations (Figures 6 and 7A). Enhanced conformational entropy and sampling of partially unfolded forms increase the rate of substrate turnover.

The chemical environment around the modification sites can regulate FAT10's impact on substrate stability. Considering Ub as a model substrate, we note that the N-terminal end of Ub is ordered, whereas the C-terminal end is disordered. FAT10 destabilizes Ub more when covalently bound at the C-terminal end. Binding at the disordered regions may allow greater conformational entropy to the FAT10 tag, giving the substrate a more significant pulling effect. Moreover, the larger conformational space of FAT10 induces more nonspecific interactions with the substrate (Figure 7E). Our results suggest that partially disordered proteins or proteins composed of disordered regions between folded domains will be more prone to destabilization by FAT10. However, further *in vitro* and cellular studies are required to understand the effect of FAT10 on partially disordered substrates.

It is interesting to investigate the specific regions in the substrate affected by FAT10 conjugation. There is a clear correlation between the NMR CSP pattern and the changes in entropy. When FAT10 is bound to the N-terminus of Ub, the CSPs are specific to β1-α1 and α2 region. The entropy increased, and Hbond stability decreased in the same regions in the substrate. When the FAT10 is bound to the C-terminal region, higher CSPs and increased entropy are observed in α2 and β5. The β1β2 turn and the C-terminal end of α1 are more entropic than the rest in free Ub[43]. The same regions in Ub were destabilized irrespective of whether FAT10 is present at the N- or C-terminus. Altogether, these results suggest that FAT10 destabilizes regions with higher entropy in the free substrate, whose magnitude depends on the modification site.

FAT10 presents an attractive system to design therapeutics by controlled protein degradation. Understanding the physical effects of FAT10 modification is essential to create a model of how the FAT10 serves as a proteasomal degradation signal and engineer it for therapeutics. We report the thermodynamic and atomistic parameters underlying the unfolding and degradation of FAT10 and its substrates. Our results suggest a loosely packed native state in FAT10 that is optimal for conjugation to substrates but unfolds instantly, allowing rapid degradation in cellular conditions. FAT10 substantially reduces substrate stability and induces partially unfolded forms in the substrate for effective proteolysis. These results reveal that the FAT10's fold and plasticity are designed to function as an efficient proteasomal degradation signal.

Conflicts of Interest

The authors declare that they have no conflicts of interest with the contents of this article.

Acknowledgments

This work was supported by the Tata Institute of Fundamental Research, Department of Atomic Energy, Government of India, under project identification no RTI 4006. The NMR data were acquired at the NCBS-TIFR NMR Facility, supported by the Department of Atomic Energy, Government of India, under

project no RTI 4006. The NMR facility is also partially supported by the Department of Biotechnology, India, under project number dbt/pr12422/med/31/287/2014. H.N. acknowledges a fellowship from UGC-CSIR, India. A.R. and P.D. acknowledge scholarships from the Department of Biotechnology, India.

Authors Contributions

H. N., A.R., and R.D. designed the study. H.N. carried out experiments and data analysis. A.R. carried out simulations and data analysis. P.D. carried out Ub denaturation experiments. S.R. carried out cell biology experiments. R.D. supervised the project. H.N. wrote the initial manuscript draft, and all authors contributed to the final draft.

References

1. Kwon, Y. T. & Ciechanover, A. The Ubiquitin Code in the Ubiquitin-Proteasome System and Autophagy. *Trends Biochem. Sci.* **42**, 873–886 (2017).

2. Peña, A. H. De, Goodall, E. A., Gates, S. N., Lander, G. C. & Martin, A. Substrate-engaged 26 S proteasome structures reveal mechanisms for ATP-hydrolysis – driven translocation. *Science (80-.).* **362**, 1018 (2018).

3. Worden, E. J., Dong, K. C. & Martin, A. An AAA Motor-Driven Mechanical Switch in Rpn11 Controls Deubiquitination at the 26S Proteasome. *Mol. Cell* **67**, 799-811.e8 (2017).

4. Wehmer, M. *et al.* Structural insights into the functional cycle of the ATPase module of the 26S proteasome. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 1305–1310 (2017).

5. Schmidtke, G., Aichem, A. & Groettrup, M. FAT10ylation as a signal for proteasomal degradation. *Biochim. Biophys. Acta - Mol. Cell Res.* **1843**, 97–102 (2014).

6. Aichem, A. *et al.* The structure of the ubiquitin-like modifier FAT10 reveals an alternative targeting mechanism for proteasomal degradation. *Nat. Commun.* **9**, 1–14 (2018).

7. Hiroi, Y. & Rechsteiner, M. Ubiquitin metabolism in HeLa cells starved of amino acids. *FEBS*

Lett. **307**, 156–161 (1992).

8. Basler, M., Buerger, S. & Groettrup, M. The ubiquitin-like modifier FAT10 in antigen processing and antimicrobial defense. *Mol. Immunol.* 10–13 (2015). doi:10.1016/j.molimm.2015.04.012

9. Raasi, S., Schmidtke, G., De Giuli, R. & Groettrup, M. A ubiquitin-like protein which is synergistically inducible by interferon-γ and tumor necrosis factor-α. *Eur. J. Immunol.* **29**, 4030–4036 (1999).

10. Hipp, M. S., Kalveram, B., Raasi, S., Groettrup, M. & Schmidtke, G. FAT10, a Ubiquitin-Independent Signal for Proteasomal Degradation. *Mol. Cell. Biol.* **25**, 3483–3491 (2005).

11. Buchsbaum, S., Bercovich, B., Ziv, T. & Ciechanover, A. Modification of the inflammatory mediator LRRFIP2 by the ubiquitin-like protein FAT10 inhibits its activity during cellular response to LPS. *Biochem. Biophys. Res. Commun.* **428**, 11–16 (2012).

12. Aichem, A. & Groettrup, M. The ubiquitin-like modifier FAT10 - much more than a proteasome-targeting signal. *J. Cell Sci.* **133**, (2020).

13. Matyskiela, M. E., Lander, G. C. & Martin, A. Conformational switching of the 26S proteasome enables substrate degradation. *Nat. Struct. Mol. Biol.* **20**, 781–788 (2013).

14. Bard, J. A. M. *et al.* The 26S Proteasome Utilizes a Kinetic Gateway to Prioritize Substrate Degradation Article. *Cell* **177**, 286-298.e15 (2019).

15. Olszewski, M. M., Williams, C., Dong, K. C. & Martin, A. The Cdc48 unfoldase prepares well-folded protein substrates for degradation by the 26S proteasome. *Commun. Biol.* **2**, (2019).

16. Carroll, E. C., Greene, E. R., Martin, A. & Marqusee, S. Site-specific ubiquitination affects protein energetics and proteasomal degradation. *Nat. Chem. Biol.* **16**, 866–875 (2020).

17. Hagai, T. & Levy, Y. Ubiquitin not only serves as a tag but also assists degradation by inducing protein unfolding. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 2001–2006 (2010).

18. Morimoto, D., Walinda, E., Fukada, H., Sugase, K. & Shirakawa, M. Ubiquitylation Directly Induces Fold Destabilization of Proteins. *Sci. Rep.* **6**, 1–9 (2016).

19. Privalov, P. L. Stability of proteins: Small Globular Proteins. in *Advances in Protein Chemistry* **33**, 167–241 (1979).

20. Park, C. & Marqusee, S. Probing the high energy states in proteins by proteolysis. *J. Mol. Biol.* **343**, 1467–1476 (2004).

21. Delaglio, F. *et al.* NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277–293 (1995).

22. Lee, W., Tonelli, M. & Markley, J. L. Structural bioinformatics NMRFAM-SPARKY : enhanced software for biomolecular NMR spectroscopy. *Bioinformatics* **31**, 1325–1327 (2015).

23. Lee, W. *et al.* I - PINE web server : an integrative probabilistic NMR assignment system for proteins. *J. Biomol. NMR* **73**, 213–222 (2019).

24. Ignatova, Z. & Gierasch, L. M. Monitoring protein stability and aggregation in vivo by real-time fluorescent labeling. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 523–528 (2004).

25. Wintrode, P. L., Makhatadze, G. I. & Privalov, P. L. Thermodynamics of ubiquitin unfolding. *Proteins* **18**, 246–253 (1994).

26. Batey, S., Nickson, A. A. & Clarke, J. Studying the folding of multidomain proteins. *HFSP J.* **2**, 365–377 (2008).

27. Lu, H. & Schulten, K. Steered molecular dynamics simulations of force-induced protein domain unfolding. *Proteins Struct. Funct. Genet.* **35**, 453–63 (1999).

28. Mücksch, C. & Urbassek, H. M. Accelerating Steered Molecular Dynamics: Toward Smaller Velocities in Forced Unfolding Simulations. *J. Chem. Theory Comput.* **12**, 1380–84 (2016).

29. Brockwell, D. J. *et al.* Mechanically unfolding the small, topologically simple protein L. *Biophys.*

J. **89**, 506–519 (2005).

30. Piana, S., Lindorff-Larsen, K. & Shaw, D. E. Atomic-level description of ubiquitin folding. *Proc. Natl. Acad. Sci.* **110**, 5915–5920 (2013).

31. Moulick, R., Das, R. & Udgaonkar, J. B. Partially unfolded forms of the prion protein populated under misfolding-promoting conditions. *J. Biol. Chem.* **290**, 25227–25240 (2015).

32. Doyle, C. M. *et al.* Concurrent Increases and Decreases in Local Stability and Conformational Heterogeneity in Cu, Zn Superoxide Dismutase Variants Revealed by Temperature-Dependence of Amide Chemical Shifts. *Biochemistry* **55**, 1346–1361 (2016).

33. Cierpicki, T. & Otlewski, J. Amide proton temperature coefficients as hydrogen bond indicators in proteins. *J. Biomol. NMR* **21**, 249–261 (2001).

34. Peters, J. H. & de Groot, B. L. Ubiquitin Dynamics in Complexes Reveal Molecular Recognition Mechanisms Beyond Induced Fit and Conformational Selection. *PLoS Comput. Biol.* **8**, (2012).

35. Zeng, D., Shen, Q. & Cho, J. H. Thermodynamic contribution of backbone conformational entropy in the binding between SH3 domain and proline-rich motif. *Biochem. Biophys. Res. Commun.* **484**, 21–26 (2017).

36. Sharp, K. A., O'Brien, E., Kasinath, V. & Wand, A. J. On the relationship between NMR-derived amide order parameters and protein backbone entropy changes. *Proteins Struct. Funct. Bioinforma.* **83**, 922–930 (2015).

37. Striebel, F., Imkamp, F., Özcelik, D. & Weber-Ban, E. Pupylation as a signal for proteasomal degradation in bacteria. *Biochim. Biophys. Acta - Mol. Cell Res.* **1843**, 103–113 (2014).

38. Ciechanover, A. & Ben-Saadon, R. N-terminal ubiquitination: More protein substrates join in. *Trends Cell Biol.* **14**, 103–106 (2004).

39. Hagai, T. & Levy, Y. Ubiquitin not only serves as a tag but also assists degradation by inducing

protein unfolding. *Proc. Natl. Acad. Sci.* **107**, 2001–2006 (2010).

40. Buchsbaum, S., Bercovich, B. & Ciechanover, a. FAT10 is a proteasomal degradation signal that is itself regulated by ubiquitination. *Mol. Biol. Cell* **23**, 225–232 (2012).

41. Khorasanizadeh, S., Peters, I. D., Roder, H. & Butt, T. R. Folding and Stability of a Tryptophan-Containing Mutant of Ubiquitin. *Biochemistry* **32**, 7054–7063 (1993).

42. Saibil, H. Chaperone machines for protein folding, unfolding and disaggregation. *Nat. Rev. Mol. Cell Biol.* **14**, 630–642 (2013).

43. Lange, O. F. *et al.* Recognition Dynamics Up to Microseconds Revealed from an RDC-Derived Ubiquitin Ensemble in Solution. *Science (80-.).* **320**, 1471–1475 (2008).

44. Dimova, N. V. *et al.* APC/C-mediated multiple monoubiquitylation provides an alternative degradation signal for cyclin B1. *Nat. Cell Biol.* **14**, 168–176 (2012).