Single-Molecule Studies on PolySUMO Proteins Reveal Their Mechanical Flexibility

Hema Chandra Kotamarthi, Riddhi Sharma, and Sri Rama Koti Ainavarapu*

Department of Chemical Sciences, Tata Institute of Fundamental Research, Colaba, Mumbai, India

ABSTRACT   Proteins with β-sandwich and β-grasp topologies are resistant to mechanical unfolding as shown by single-molecule force spectroscopy studies. Their high mechanical stability has generally been associated with the mechanical clamp geometry present at the termini. However, there is also evidence for the importance of interactions other than the mechanical clamp in providing mechanical stability, which needs to be tested thoroughly. Here, we report the mechanical unfolding properties of ubiquitin-like proteins (SUMO1 and SUMO2) and their comparison with those of ubiquitin. Although ubiquitin and SUMOs have similar size and structural topology, they differ in their sequences and structural contacts, making them ideal candidates to understand the variations in the mechanical stability of a given protein topology. We observe a two-state unfolding pathway for SUMO1 and SUMO2, similar to that of ubiquitin. Nevertheless, the unfolding forces of SUMO1 (~130 pN) and SUMO2 (~120 pN) are lower than that of ubiquitin (~190 pN) at a pulling speed of 400 nm/s, indicating their lower mechanical stability. The mechanical stabilities of SUMO proteins and ubiquitin are well correlated with the number of interresidue contacts present in their structures. From pulling speed-dependent mechanical unfolding experiments and Monte Carlo simulations, we find that the unfolding potential widths of SUMO1 (~0.51 nm) and SUMO2 (~0.33 nm) are much larger than that of ubiquitin (~0.19 nm), indicating that SUMO1 is six times and SUMO2 is three times mechanically more flexible than ubiquitin. These findings might also be important in understanding the functional differences between ubiquitin and SUMOs.

INTRODUCTION

For more than a decade, it has been shown that pulling of polyproteins using single-molecule force spectroscopy (SMFS) to be very useful in characterizing the mechanical properties and the unfolding pathways of proteins (1–10). Proteins with β-sandwich topology or β-grasp topologies received much attention for their mechanical resistance against unfolding in single-molecule studies. In these studies, the mechanical stability has generally been attributed to the geometry in which the N- and C-terminal β-strands are parallel and directly connected by H-bonding (mechanical clamp). Mechanical clamp geometry is considered to provide resistance and rupture of it leads to unfolding as seen in single-molecule studies (11,12). On the other hand, it has been observed that the other interactions in the proteins, which are away from the mechanical clamp, are also important in providing mechanical resistance. In these studies, proteins with the same topology were used to provide support for the significance of interactions away from the mechanical clamp in providing the mechanical stability (13,14). Although, these previous studies provide vital evidence on the role of various interactions in proteins, the proteins were not of the same size to have a direct comparison between the mechanical stability and interactions. Here, we addressed this issue by comparing the mechanical properties of ubiquitin and small ubiquitin-related modifiers (SUMOs), which have the same topology, and more importantly, the same size.

SUMOs are members of the ubiquitin-like protein family that act as reversible posttranslational modifiers of proteins in eukaryotes (15–17). The mechanism of SUMO ligation to target proteins, which is called SUMOylation, is similar to that of ubiquitination. SUMOylation occurs through the formation of an isopeptide bond between the C-terminal Gly residue of SUMO protein and ε-amino group of Lys of the target protein through the action of E1, E2, and E3 enzymes in a manner similar to that of ubiquitination. Ubiquitin's function is well defined in targeting proteins for degradation by tagging them with polyubiquitin. Furthermore, polyubiquitin chains are shown to resist mechanical unfolding in single-molecule experiments, and it has been associated with their function in proteasomal degradation (18). Unlike ubiquitin, the function of SUMOs is to modulate the function of the target protein. SUMO targets are involved in various cellular processes such as regulating mitochondrial function, chromatin structure, and signal transduction. In all these processes, SUMO proteins bind to other proteins and modify them (15,16,19,20).

In humans, there are four different SUMO proteins: SUMO1, SUMO2, SUMO3, and SUMO4. Although the sequences of SUMO2 and SUMO3 are >95% identical, they differ widely from that of SUMO1 with only ~50% identity. Here, the sequence identity between two proteins is a pairwise identity between their amino acid sequences. The sequence identity of SUMO1 and SUMO2 with that of ubiquitin is below 20% (Fig. 1 A). Nevertheless, SUMO1, SUMO2, and SUMO3 share high structural homology (>95%) with ubiquitin (21–24). The structural homology between two protein structures was calculated using
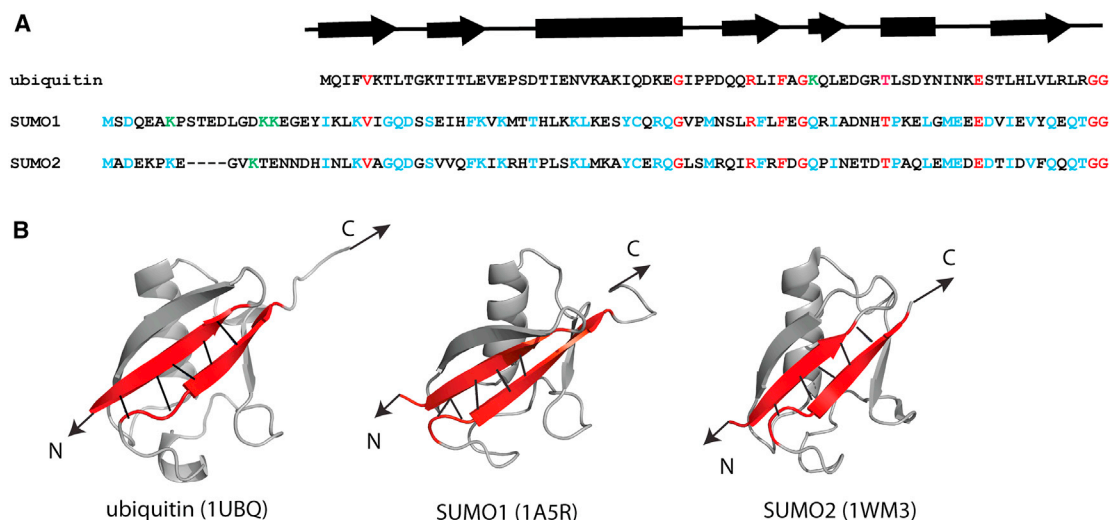
FIGURE 1 (A) The amino acid sequences of ubiquitin, SUMO1, and SUMO2. The residues that are identical in all three proteins are highlighted in red. In SUMO1 and SUMO2 there exists a long structureless sequence (~20 residues) at N-terminus, which is absent in ubiquitin. The sequence identity between SUMO1/ubiquitin, SUMO2/ubiquitin pairs is ~16%. The sequence identity between the SUMO1/SUMO2 pair is high (~50%) and the residues identical in both of them are highlighted in blue. The C-terminal diglycine sequence is preserved in all of them. Ubiquitin forms Lys-48-C-linked polyubiquitin, whereas SUMO2 forms Lys-11-C-linked polymer chains. SUMO1 forms polymer chains in vitro through Lys-7, Lys-16, and Lys-17, which are highlighted in green. The secondary structural elements of ubiquitin are indicated directly above the sequence. (B) Structures of SUMO1, SUMO2, and ubiquitin. PDB IDs are given in parenthesis. The N-terminus structureless regions in SUMO1 and SUMO2 are not shown. In all the structures, the terminal β-strands (shown in *red*) are parallel to each other and connected by five H-bonds between the backbones (*black lines*). The arrows indicate the pulling direction used in the mechanical unfolding experiments.

TopMatch protein structure alignment software (25,26). SUMOs have a characteristic long unstructured N-terminus that is absent in ubiquitin and other ubiquitin-like proteins (16,24). Structures of SUMOs consist of a classical β-grasp ubiquitin fold with four β-strands wrapped around a α-helix (Fig. 1 B). Despite their high structural similarity, SUMOs and ubiquitin are functionally divergent. This also raises an important question on what structural nuances and dynamics of these proteins make them functionally different. Furthermore, the terminal β-strands in the ubiquitin fold are directly connected to each other via backbone H-bonding, which acts as a mechanical clamp during unfolding as shown earlier by Carrion-Vazquez et al. (18). Based on the structural similarity between SUMOs and ubiquitin, their mechanical properties are expected to be similar, however, this remains to be tested experimentally. It must be noted that SUMO proteins are also found to undergo polymerization in a manner similar to polyubiquitination and it is not completely understood why such polySUMO proteins exist in vivo (27,28). The polymerization of SUMO proteins could be carried out in vitro as well (29). Hence, it would be important to study the biophysical properties of SUMOs and compare them with ubiquitin to gain further understanding on the functional diversity and the possible role of SUMO polymers in vivo.

Here, we report the mechanical properties of SUMO1 and SUMO2 measured in constant velocity pulling experiments using SMFS and compare them with those of ubiquitin. We find that SUMO1 and SUMO2 unfold at a much lower force than that of ubiquitin despite having similar structure. Furthermore, we have also varied pulling speed, used Monte Carlo simulations, and the Bell-like model to obtain the details of their unfolding energy landscapes. We find differences in their underlying energy landscapes, in terms of the distance to the transition state from the native state.

MATERIALS AND METHODS

Polyprotein engineering

The genes of human SUMO1 and human SUMO2 were modified to insert the restriction sites of BamHI at the 5′ end and BglII, KpnI at the 3′ end. The genes were then cloned into pQE80L vector between BamHI and KpnI sites. The genes of (SUMO1)$_8$ and (SUMO2)$_8$ were constructed by an iterative cloning method as described by Carrion-Vazquez et al. (2). The octameric proteins, (SUMO1)$_8$ and (SUMO2)$_8$, were overexpressed in the BLR DE3 strain of *Escherichia coli* by inducing with 1 mM IPTG (isopropylthio-β-D-1-thiogalactopyranoside) for 6 h after the OD$_{600}$ of the cell culture has reached 0.6. The harvested cells were suspended in phosphate buffered saline (PBS) (pH 7.4) containing 1 mM DTT (dithiothreitol) and 0.1 mM PMSF (phenylmethanesulfonyl fluoride). The cells were lysed by sonication, centrifuged, and the supernatant was applied to a column of Ni-NTA-coated agarose beads. The beads were washed with PBS containing 20 mM imidazole and the proteins were eluted with PBS containing 250 mM imidazole at pH 7.4. They were further purified by size exclusion chromatography using a superdex200 column (Amersham Biosciences). The purity of proteins was checked on sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) gel. All purified proteins were stored at 4°C. Polyubiquitin was expressed and purified using protocols described earlier by Carrion-Vazquez et al. (18).

SMFS

Single-molecule pulling experiments were performed on a custom-built atomic force microscope as described elsewhere (10). Approximately, 5 μl of polyprotein sample (~2 μM) was added to ~50 μl of PBS (pH 7.4) on a gold-coated glass coverslip. Gold-coated reflective cantilevers with a silicon-nitride tip with spring constants ~35 pN/nm were purchased from Veeco, Singapore. Calibration of cantilevers was done using the equipartition theorem before each pulling experiment (30). All the experiments were performed at room temperature. All force-extension (FX) traces were collected at a sampling rate of 5 kHz.

Data analysis

In the data analysis, force peaks in FX traces were fitted to a worm-like chain (WLC) model of polymer elasticity (31) using Eq. 1

$$F(x) = \frac{k_B T}{p}\left[\frac{1}{4}\left(1 - \frac{x}{L_c}\right)^{-2} - \frac{1}{4} + \frac{x}{L_c}\right], \qquad (1)$$

where p, L_c denotes the persistence length and contour length, respectively, k_B is Boltzmann's constant and T is absolute temperature. The persistence length was varied in the range of 0.3–0.8 nm with an average value of ~0.5 nm. This range of persistence length for the WLC model was found to best fit the FX traces.

Spontaneous unfolding rate constant (k_u^0) and distance to the unfolding transition state (Δx_u)

The spontaneous unfolding rate constant (k_u^0) and the distance to the unfolding transition state (Δx_u) were calculated using two different methods: Monte Carlo simulations and by fitting the experimental data to the Bell-like model. The k_u^0, Δx_u were calculated using Monte Carlo simulations as described by Oberhauser et al. (32). In brief, the dependence of unfolding force on pulling speed was modeled assuming a two-state Markov process (1,33). Initially, all the domains in a polyprotein chain were assumed to be folded (N_f) and were pulled at a given speed and the force was calculated using the WLC model as described previously. The probability of unfolding (P_u) of any of the domains is given by $P_u = N_f k(F)\Delta t$, where $k(F) = k_u^0 \cdot e^{F\Delta x_u/k_B T}$. The unfolding event and force at which it occurred was recorded when the probability P_u, was greater than a random number between 0 and 1. The procedure was repeated until all the domains in the polyprotein chain unfolded and the unfolding force histograms were made. The values of k_u^0 and Δx_u were varied such that the unfolding force histograms from simulations matched with those obtained experimentally at different pulling speeds.

We have also fitted the pulling-speed-dependent experimental unfolding forces, $F(v)$, to Bell-Evans-Ritchie approximation (34–37) and extracted the k_u^0, Δx_u values. The following equation was used to fit the experimental values.

$$F(v) = \frac{k_B T}{\Delta x_u}\ln\left(\frac{v\,\Delta x_u}{k_u^0\,k_B\,T}\right), \qquad (2)$$

where v is the loading rate.

Transition state energy barrier and spring constant of the unfolding potential

The transition state energy barrier and spring constant of the unfolding potential were calculated using the method suggested by Dietz et al. (38). The transition state barrier height ($\Delta G^{\#}$) was calculated using the Arrhenius equation $\Delta G^{\#} = -k_B T \times \ln(k_u^0/k_A)$, where k_B is Boltzmann's constant, T is temperature (= 298 K), k_u^0 is the rate constant for spontaneous unfolding and k_A is the Arrhenius frequency factor. Here, k_A is taken as 10^9s^{-1} (38,39). The unfolding potential is assumed to be harmonic and its spring constant (k_s) is calculated using the equation $k_s = 2\,\Delta G^{\#}/(\Delta x_u)^2$, where Δx_u is the distance to the transition state or width of the unfolding potential.

RESULTS AND DISCUSSION

SUMO1 and SUMO2 unfold at a lower mechanical force than that of ubiquitin

We performed pulling experiments on polyproteins (ubiquitin)$_9$, (SUMO1)$_8$, and (SUMO2)$_8$ using SMFS at a pulling speed of 400 nm/s. The details of the protein engineering and the experimental technique are given in the Materials and Methods section. A typical FX trace of each polyprotein is shown in Fig. 2 A. For all three cases, the FX traces show a sawtooth pattern of force peaks indicating the sequential unfolding of individual protein units in the polyprotein. Furthermore, the spacing of ~24 nm between adjacent peaks indicates that all three proteins have approximately the same number of amino acids contributing to the change in contour length upon unfolding. In these FX traces, the last force peak at >400 pN is due to the detachment of the protein either from the tip or from the substrate. For further analysis, we have taken all the FX traces that have at least four unfolding force peaks. The single-molecule FX traces were fitted to the WLC model (31) and the change in contour length of protein upon mechanical unfolding is ~24 nm in all three cases. The histograms of the change in contour length upon unfolding for all three proteins are given in the Supporting Material (Fig. S1). However, their unfolding forces are different: ubiquitin (~190 pN) unfolds at a higher mechanical force than SUMO1 (~130 pN) and SUMO2 (~120 pN) at a pulling speed of 400 nm/s. The unfolding force histograms are given in Fig. 2 B. Our results on ubiquitin are in agreement with those reported earlier (18,40). The results are given in Table 1.

SUMO1 and SUMO2 have 97 and 93 residues, respectively (Fig. 1 A). SUMO proteins are known to have a long and highly flexible N-terminus, which is absent in ubiquitin (24). A contiguous stretch of 76 residues spanning to the C-terminus in SUMO1 and SUMO2, forms a structure similar to that of ubiquitin with >95% homology (Fig. 1 B). Based on the high similarity between the contour length change and the structures of SUMOs and ubiquitin, it can be envisaged that the structureless region at the N-terminus of SUMO1 and SUMO2 unravels without any discernible force peak in the FX trace and the contiguous stretch of 76 residues spanning to the C-terminus unravel in a cooperative manner resulting in a single force peak with a contour length change of 24 nm (Fig. 2 A). Indeed, we have analyzed the featureless spacer before the sawtooth pattern in FX traces for ubiquitin, SUMO1, and SUMO2 (Fig. S2). The spacer lengths are longer for SUMO1 and SUMO2 in
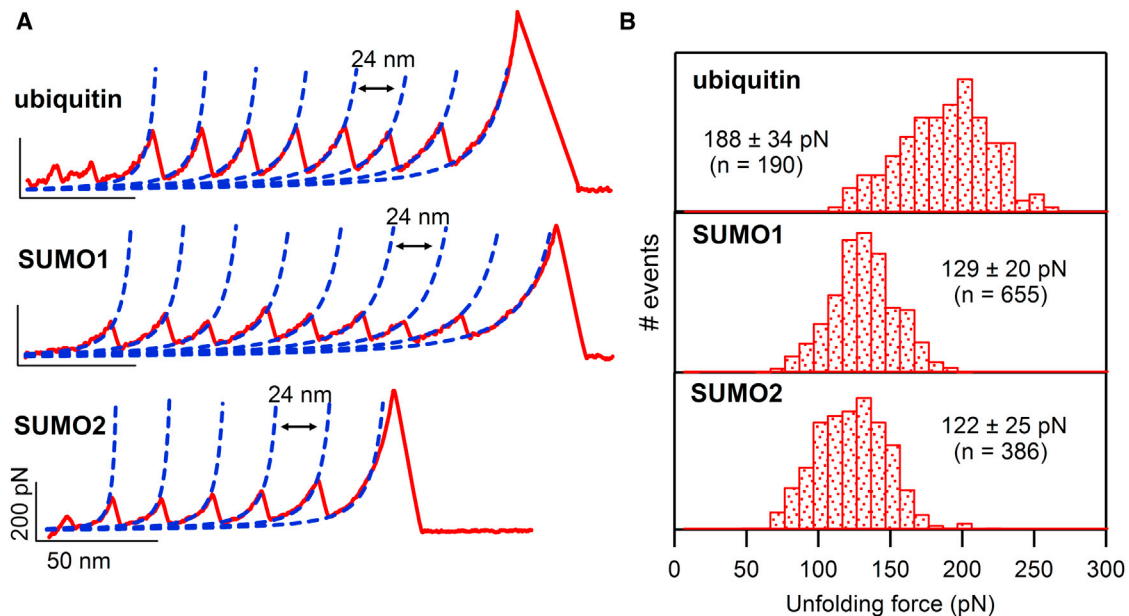
FIGURE 2 SUMOs are mechanically weaker compared to ubiquitin. (*A*) (*top to bottom*) Representative single-molecule force-extension (FX) traces of (ubiquitin)$_9$, (SUMO1)$_8$, and (SUMO2)$_8$ at a pulling speed of 400 nm/s showing the sawtooth pattern with a contour length increment ~24 nm. The traces were fit with the WLC model (*dashed lines*). The scale bars are 200 pN (*vertical*) and 50 nm (*horizontal*). (*B*) Unfolding force histograms (see Table 1 for more details).

comparison to ubiquitin, confirming that the structureless region at the N-terminus of SUMO1 and SUMO2 unravels without any mechanical resistance. As expected, the change in contour lengths of SUMO1 and SUMO2 are exactly identical to that of ubiquitin (Fig. 2, Fig. S1, and Table 1).

Role of noncovalent interactions on protein mechanical stability

In the case of ubiquitin, the origin of mechanical stability was attributed to the five H-bonds between the force-bearing parallel β-strands at the termini (18). Terminal β-strands of SUMO1 and SUMO2 are also connected by five H-bonds (Fig. 1 *B*). Despite the same pulling geometry and rupture of the same number of H-bonds between the force-bearing structures, the unfolding force is very low for SUMO1 and SUMO2.

Previously, the parallel geometry of the terminal β-strands and the number of H-bonds connecting them in

proteins is hypothesized to be the main reason for their very high mechanical stability (11) and their rupture leads to the mechanical unfolding as shown for many β-sheet and α/β proteins like I27, GB1, etc. (41,42). Proteins of this structural geometry at the termini have been found to unfold at higher forces (>180 pN at a pulling speed of ~400 nm/s). However, proteins that lack this structural geometry at the termini, such as barnase, barstar, and DHFR tend to be mechanically compliant and unfold at much lower unfolding forces (<100 pN at a pulling speed of ~400 nm/s) (43–45). Based on these general observations, SUMO1 and SUMO2 are also expected to be as strong as ubiquitin. However, our experiments demonstrate that this is not the case and SUMO1 and SUMO2 are mechanically weaker compared to ubiquitin.

In previous experiments, Li et al. (46) have shown that I1 and I27, the 1st and the 27th domains of I-band of human cardiac titin, respectively, have different mechanical stabilities as well as different distances to transition states

TABLE 1 Mechanical properties of ubiquitin and SUMO proteins

Protein	Change in contour length (nm)*,†	Unfolding force (pN)‡	Δx_u(nm)		k_u^0(s^{-1})		$\Delta G^{\#}$ ($k_B T$)	k_s(N/m)
Ubiquitin	23.9 ± 1.2 ($n = 190$)	188 ± 2	0.19§	0.15 ± 0.01¶	$8 \times 10^{-3\S}$	0.45 ± 0.12¶	25.5	5.8
SUMO1	24.1 ± 0.9 ($n = 655$)	129 ± 0.8	0.51§	0.51 ± 0.04¶	$1.15 \times 10^{-5\S}$	$1.8 \times 10^{-4} \pm 1.6 \times 10^{-4}$¶	32.1	1.0
SUMO2	24.1 ± 1.3 ($n = 386$)	122 ± 1	0.33§	0.28 ± 0.02¶	$5 \times 10^{-3\S}$	0.32 ± 0.19¶	26.1	2.0

*Mean ± SD.

†The n in the parentheses is the number of events used in the data analysis.

‡Mean ± SE.

§From Monte Carlo simulation.

¶From Bell-Evans-Ritchie approximation.

despite similar structures. The I1 has a much lower mechanical unfolding force (130 pN) than that of I27 (210 pN) at the same pulling speed. These differences in mechanical properties were attributed to the variation in the distribution of H-bonds between the two terminal β-strands, hydrophobic interactions, and salt bridges in the protein (46–48). From our experiments, we found out that the mechanical properties of SUMO1, SUMO2, and ubiquitin have similar features as that of I1 and I27. However, the distinctly different mechanical properties of SUMOs from that of ubiquitin cannot be due to the difference in the number of H-bonds joining the two terminal β-strands as this is the same (five) in all the cases (Fig. 1 B). Hence, it is very likely that the mechanical stability of SUMOs is not solely dependent on the H-bonding at the termini but is also influenced by other interactions within the protein. This issue was earlier highlighted by research groups of Clarke (14,44) and Radford (13,49) where they have used interresidue contact maps in interpreting their experimental findings. TNfn3, a fibronectin-like domain from tenascin, despite having the same topology as I27, unfolds at a much lower force (~100 pN) and the resistance to mechanical unfolding of TNfn3 was attributed to interresidue interactions present in the protein core in addition to those between the termini (14). Similarly, while comparing protein L with ubiquitin, Brockwell et al. (13) suggested that hydrophobic interactions in the protein structure might play an important role in modulating the mechanical stability of protein L because it was found to unfold at a lower force (~130 pN) than ubiquitin at a similar pulling speed despite having the same structural fold.

To understand the discrepancy in the mechanical stability of SUMO1 and SUMO2 with that of ubiquitin, we have adopted a similar strategy of analyzing interresidue con-

tacts in these proteins as discussed previously. In Fig. 3 A, the interresidue side-chain contact maps of SUMO1 and SUMO2 in comparison to ubiquitin are shown. It clearly shows that the total number of contacts are different for SUMO1 (138 contacts), SUMO2 (137 contacts), and ubiquitin (154 contacts) as calculated using CM-view software (50). In this calculation, an interresidue contact was assumed to be present if the shortest distance between side-chain atoms of different residues is <5 Å. We have also varied the cutoff distance and found that ubiquitin always has a higher number of contacts than SUMO1 and SUMO2 (Fig. S3). We have also calculated the number of side-chain contacts between the N-terminal ($\beta\beta\alpha$) and C-terminal ($\beta\beta$) halves, which form the force bearing regions of each protein when pulled in the N-C direction (Fig. S4). The number of contacts between the halves also follow the same trend as the overall contacts: ubiquitin (45 contacts) > SUMO1 (37 contacts) > SUMO2 (34 contacts). We have dissected the total number of side-chain contacts into hydrophilic/H-bond, aromatic, and hydrophobic interactions using CSU software (51). A detailed classification of interactions for each protein is provided in the Supporting Material, Table S1. Nearly 65% interactions are hydrophobic and 35% hydrophilic/H-bond interactions for all three proteins with a couple of aromatic interactions in the case of SUMO1 and SUMO2. The difference in mechanical stability between SUMOs and ubiquitin does not seem to be due to the hydrophilic or hydrophobic interactions alone as both are higher in ubiquitin. Hence, it is possible that an overall lower number of contacts present in SUMOs might be the reason for its lower mechanical stability. Using CSU software, we have also estimated the contact surface area of these contacts and it follows the same trend as the number of contacts (Table S1). Contact maps were used earlier
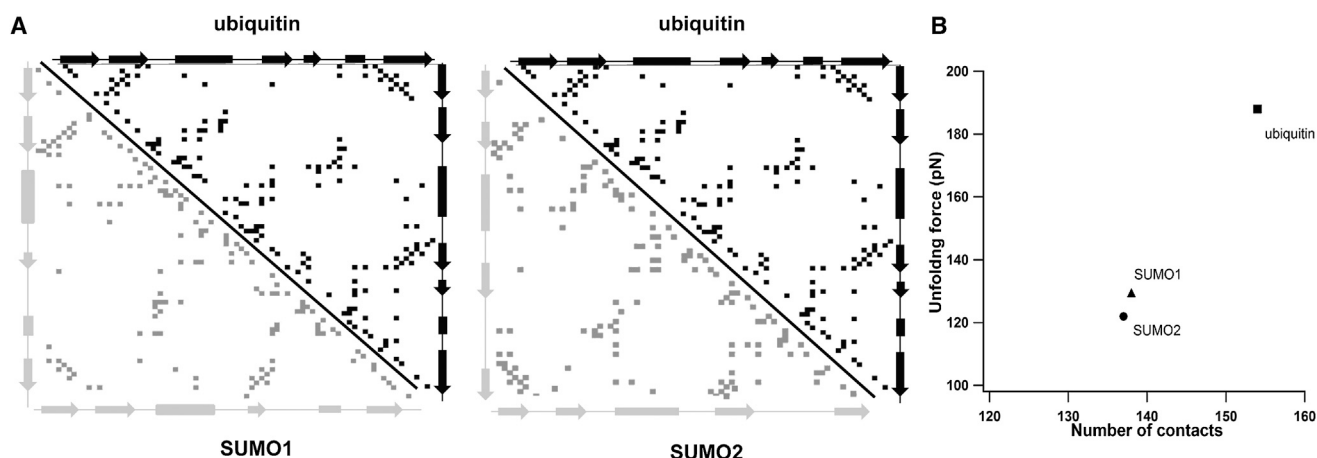

FIGURE 3   The number of interresidue contacts correlates with the unfolding force. (A) Contact maps of ubiquitin, SUMO1, and SUMO2. Ubiquitin paired with SUMO1 and SUMO2 separately for direct comparison. The definition of an interresidue contact is defined in the text. The total number of contacts in ubiquitin (154) are higher than SUMO1 (138) and SUMO2 (137), and this could be a reason for the lower mechanical stability of SUMO1 and SUMO2 compared to ubiquitin. (B) Plot of number of contacts versus unfolding forces for ubiquitin (■), SUMO1 (▲), and SUMO2 (●).

to explain the differences in the mechanical stabilities of protein L and ubiquitin (13). However, the sizes of protein L (62 residues) and ubiquitin (76 residues) are different, though their overall fold is the same (13). Here, we have compared proteins from the same subfamily, with similar structural fold, and more important, having the same number of amino acids. There is a good correlation between the number of contacts and the unfolding force for SUMO1, SUMO2, and ubiquitin (Fig. 3 B). The contact maps seem to explain the discrepancies in the mechanical stabilities. Furthermore, we also investigated the details of the unfolding energy landscape of these proteins by performing pulling speed-dependent mechanical unfolding experiments.

Unfolding energy landscapes of SUMO1, SUMO2, and ubiquitin

A semilogarithmic plot of unfolding force versus pulling speed for SUMO1, SUMO2, and ubiquitin is shown in Fig. 4. The speed-dependent unfolding force histograms are given in the Supporting Material (Fig. S5). The unfolding forces of SUMO1 and SUMO2 are consistently below ubiquitin throughout the pulling speed range (Fig. 4). Similar to Fig. 3 B, we have compared the number of contacts with unfolding forces at different pulling speeds (Fig. S6). For SUMOs and ubiquitin in general, the relationship that lower the number of contacts mean lower the mechanical stability seems to be valid in the entire pulling speed range (40–4000 nm/s). From this data, it can be said that the number of contacts could serve as an indicator of mechanical strength when comparisons have to be made within a class of structurally homologous proteins.
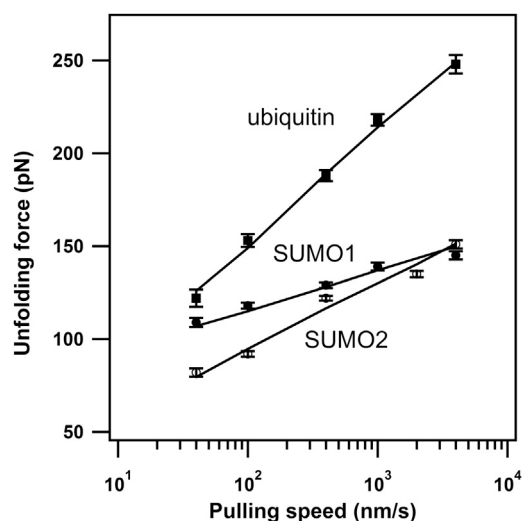

FIGURE 4 Pulling speed-dependent unfolding properties of ubiquitin, SUMO1, and SUMO2. Plot of unfolding force versus pulling speed. Error bars indicate SE. Solid lines are fits from Monte Carlo simulations. The results of simulations are given in Table 1.

In Fig. 4, the slopes of the experimental data are different for the three proteins indicating the differences in their unfolding potentials. To extract the unfolding energy landscape parameters, we have performed Monte Carlo simulations assuming a two-state unfolding process where the native state and unfolded states are separated by a single transition state barrier (32). The simulated data along with the experimental data are shown in Fig. 4. We found out that the distance to the unfolding transition state (Δx_u) to be 0.19 nm for ubiquitin, 0.51 nm for SUMO1, and 0.33 nm for SUMO2. Our measured Δx_u for ubiquitin is comparable to the earlier reported values, 0.225 nm from constant velocity measurements by Carrion-Vazquez et al. (18) and 0.17 nm from constant force measurements by Schlierf et al. (52). The calculated spontaneous unfolding rates (k_u^0) for three proteins are given in Table 1. To independently estimate Δx_u and k_u^0, we have used Bell-Evans-Ritchie approximation as described previously (34–36). The estimates from this model are also given in Table 1. The Δx_u values from this model are in excellent agreement with that of Monte Carlo simulations. The k_u^0 values from this model differ from Monte Carlo simulations by about an order of magnitude. Both Monte Carlo simulations and the Bell-like model give more accurate estimates of Δx_u than k_u^0. In fact, the Δx_u values from both the simulation and the model estimate that the unfolding potential of ubiquitin is much narrower than SUMO1 and SUMO2. These results can be explained in the form of an energy landscape where the activation energies are comparable, but the distance to the transition state (Δx_u) is longer for SUMO1 and SUMO2 than ubiquitin and hence a lower unfolding force applied over a longer distance can decrease the barrier for SUMO1 and SUMO2 during the mechanical unfolding. A schematic of the unfolding energy landscapes of SUMO1, SUMO2 and ubiquitin describing the experimental and simulation results is shown in Fig. 5.

As seen in the previous section, the number of contacts is directly correlated with mechanical stability, whereas it is not correlated with k_u^0 and Δx_u. Mechanical stability of a protein is both due to the unfolding barrier (or k_u^0) and Δx_u, but either parameter alone might not explain it. Mechanical resistance can be correlated to an unfolding barrier for proteins that have the same Δx_u and vice versa. However, in the case of SUMO1, SUMO2, and ubiquitin, both k_u^0 and Δx_u are varying, and hence mechanical resistance cannot be directly correlated to k_u^0 (or Δx_u) alone. Similarly, it is possible that the number of contacts that impart stability to a protein might not be correlated independently to the unfolding barrier or Δx_u.

Furthermore, the Δx_u is the magnitude of deformation along the pulling direction that occurs in protein to reach the transition state before crossing the unfolding energy barrier and it is usually taken as a measure of protein deformation response (38). We have estimated the stiffness of ubiquitin, SUMO1, and SUMO2 from their potential width
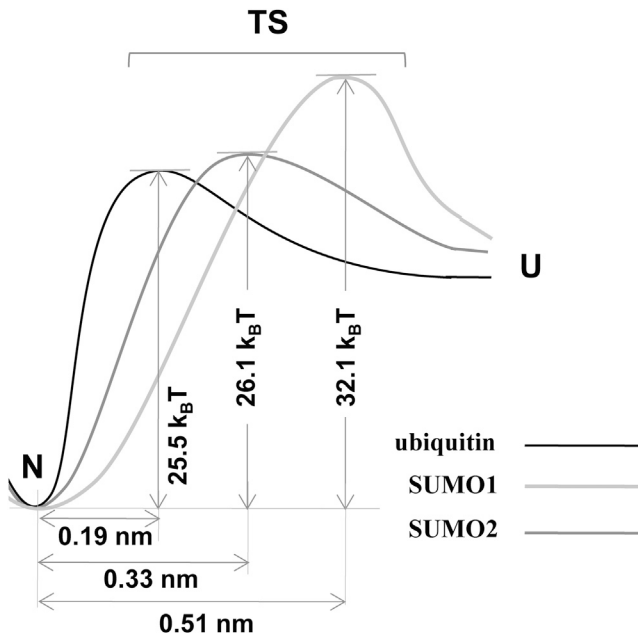
FIGURE 5 A schematic of the unfolding energy landscape. The unfolding potential width Δx_u (or the distance to the transition state (TS)) of ubiquitin is much shorter than SUMO1 and SUMO2. See text and Table 1 for more details.

(Δx_u) and the energy barrier as suggested by Dietz et al. (38). We have used the Δx_u and k_u^0 values obtained from Monte Carlo simulations in estimating the protein stiffness. We found out that ubiquitin (5.8 N/m) has a much larger spring constant for the unfolding potential (k_s) than SUMO1 (1.0 N/m) and SUMO2 (2.0 N/m), which was calculated from the curvature of the potential well that precedes the unfolding energy barrier. This suggests that SUMO1 is six times and SUMO2 is three times more flexible than ubiquitin during mechanical unfolding. Between SUMO1 and SUMO2, SUMO2 is more similar to ubiquitin based on the unfolding energy landscape parameters—potential width, barrier height, and stiffness (Fig. 5 and Table 1). This is despite SUMO2 being less similar to ubiquitin (14% sequence identity) than SUMO1 (46% sequence identity) (Fig. 1). This could be explained by comparing the structures of SUMOs and ubiquitin, as done by Ding et al. (23), where it was shown that the root mean-square deviation value of average core structure of SUMO1 and SUMO3 pair is 2.2 Å, whereas the root mean-square deviation value for SUMO3, ubiquitin pair is 1.4 Å, which means the backbone core structure of SUMO3 is closer to ubiquitin than SUMO1. SUMO3 can represent SUMO2 as there is 97% sequence identity between them and the residues that are not common lie only in the unstructured N-terminus of SUMO2. On the basis of this comparison, it can be said that the backbone core structure of SUMO2 is more similar to ubiquitin than SUMO1. Hence, it is expected that the SUMO2/ubiquitin pair will have similar unfolding properties than the SUMO1/

ubiquitin pair. Indeed, this is in concurrence with our observations (Fig. 5).

Biological significance of mechanical properties of SUMOs

Our finding that SUMO1 and SUMO2 are mechanically weaker and flexible when compared to ubiquitin, might have implications for their function. In addition, the polyproteins (SUMO1)$_8$ and (SUMO2)$_8$ studied here are analogous to multichain SUMO polymers found in vitro and in vivo. SUMO2 and SUMO3 are known to form multimers via an internal consensus site (Val-Lys-Thr-Glu) present at the N-terminus with the C-terminal Gly and can form Lys-11-C-linked chains (53,54). This internal consensus site for SUMOylation is absent in SUMO1. Interestingly, SUMO1 is found to form polymers linked via Lys-7, Lys-16, and Lys-17 in vitro (54,55). SUMO1 by itself does not form polymers in vivo but participates as a chain terminator in mixed SUMOylation with SUMO2 and SUMO3. However, in all SUMOs, the polymerization sites are located within the structureless N-terminal extensions that are absent in ubiquitin. These polySUMOs are equivalent to polyproteins (SUMO1)$_8$ and (SUMO2)$_8$ used in the current study, as SUMO proteins are linked at the N- and C-termini in a head-to-tail manner via peptide bonds. Interestingly, the mechanical stability of polySUMOs and N-C-linked polyubiquitin is higher than the physiologically relevant Lys-48-C-linked polyubiquitin (18). The differences in the unfolding forces are due to the anisotropic nature of the force, because these proteins are pulled along different directions. However, more studies would be required to understand if the higher mechanical resistance of polySUMOs has any role in proteasomal degradation given the recent findings on the cross talk between polySUMOs and polyubiquitin (15,56).

CONCLUSION

In summary, we have constructed octameric polyproteins of SUMO1 and SUMO2, and measured their mechanical properties. Both SUMO1 and SUMO2 have lower unfolding forces than ubiquitin. Although, SUMO1, SUMO2, and ubiquitin have the same overall topology and the structural arrangement at the termini (terminal β-strands, which are H-bonded to each other), their structures have a different number of contacts. The measured mechanical stabilities are in correlation with their interresidue contacts. Our study further supports the hypothesis that the interresidue contacts, in addition to mechanical clamp, are important in providing proteins with mechanical resistance and stability. Our study is also important because it used proteins not only from the same structural family but also of the same size. Furthermore, the unfolding energy landscape parameters obtained from experiments, simulations, and theoretical

models suggest that SUMO1 and SUMO2 are mechanically more flexible than ubiquitin.

SUPPORTING MATERIAL

REFERENCES

1. Rief, M., M. Gautel, …, H. E. Gaub. 1997. Reversible unfolding of individual titin immunoglobulin domains by AFM. *Science.* 276:1109–1112.

2. Carrion-Vazquez, M., A. F. Oberhauser, …, J. M. Fernandez. 1999. Mechanical and chemical unfolding of a single protein: a comparison. *Proc. Natl. Acad. Sci. USA.* 96:3694–3699.

3. Carrion-Vazquez, M., A. F. Oberhauser, …, J. M. Fernandez. 2000. Mechanical design of proteins studied by single-molecule force spectroscopy and protein engineering. *Prog. Biophys. Mol. Biol.* 74:63–91.

4. Crampton, N., and D. J. Brockwell. 2010. Unravelling the design principles for single protein mechanical strength. *Curr. Opin. Struct. Biol.* 20:508–517.

5. Li, H., and Y. Cao. 2010. Protein mechanics: from single molecules to functional biomaterials. *Acc. Chem. Res.* 43:1331–1341.

6. Puchner, E. M., and H. E. Gaub. 2009. Force and function: probing proteins with AFM-based force spectroscopy. *Curr. Opin. Struct. Biol.* 19:605–614.

7. Galera-Prat, A., A. Gómez-Sicilia, …, M. Carrión-Vázquez. 2010. Understanding biology by stretching proteins: recent progress. *Curr. Opin. Struct. Biol.* 20:63–69.

8. Ainavarapu, S. R., A. P. Wiita, …, J. M. Fernandez. 2008. A single-molecule assay to directly identify solvent-accessible disulfide bonds and probe their effect on protein folding. *J. Am. Chem. Soc.* 130:436–437.

9. Bornschlögl, T., and M. Rief. 2011. Single-molecule protein unfolding and refolding using atomic force microscopy. *Methods Mol. Biol.* 783:233–250.

10. Aggarwal, V., S. R. Kulothungan, …, S. R. Ainavarapu. 2011. Ligand-modulated parallel mechanical unfolding pathways of maltose-binding proteins. *J. Biol. Chem.* 286:28056–28065.

11. Eom, K., P. C. Li, …, G. J. Rodin. 2003. Relationship between the mechanical properties and topology of cross-linked polymer molecules: parallel strands maximize the strength of model polymers and protein domains. *J. Phys. Chem. B.* 107:8730–8733.

12. Hoffmann, T., and L. Dougan. 2012. Single molecule force spectroscopy using polyproteins. *Chem. Soc. Rev.* 41:4781–4796.

13. Brockwell, D. J., G. S. Beddard, …, S. E. Radford. 2005. Mechanically unfolding the small, topologically simple protein L. *Biophys. J.* 89:506–519.

14. Ng, S. P., R. W. Rounsevell, …, J. Clarke. 2005. Mechanical unfolding of TNfn3: the unfolding pathway of a fnIII domain probed by protein engineering, AFM and MD simulation. *J. Mol. Biol.* 350:776–789.

15. Hay, R. T. 2005. SUMO: a history of modification. *Mol. Cell.* 18:1–12.

16. Geiss-Friedlander, R., and F. Melchior. 2007. Concepts in sumoylation: a decade on. *Nat. Rev. Mol. Cell Biol.* 8:947–956.

17. Vertegaal, A. C. 2011. Uncovering ubiquitin and ubiquitin-like signaling networks. *Chem. Rev.* 111:7923–7940.

18. Carrion-Vazquez, M., H. Li, …, J. M. Fernandez. 2003. The mechanical stability of ubiquitin is linkage dependent. *Nat. Struct. Biol.* 10:738–743.

19. Müller, S., C. Hoege, …, S. Jentsch. 2001. SUMO, ubiquitin's mysterious cousin. *Nat. Rev. Mol. Cell Biol.* 2:202–210.

20. Melchior, F. 2000. SUMO—nonclassical ubiquitin. *Annu. Rev. Cell Dev. Biol.* 16:591–626.

21. Song, J., Z. Zhang, …, Y. Chen. 2005. Small ubiquitin-like modifier (SUMO) recognition of a SUMO binding motif: a reversal of the bound orientation. *J. Biol. Chem.* 280:40122–40129.

22. Huang, W. C., T. P. Ko, …, A. H. Wang. 2004. Crystal structures of the human SUMO-2 protein at 1.6 A and 1.2 A resolution: implication on the functional differences of SUMO proteins. *Eur. J. Biochem.* 271:4114–4122.

23. Ding, H., Y. Xu, …, Y. Shi. 2005. Solution structure of human SUMO-3 C47S and its binding surface for Ubc9. *Biochemistry.* 44:2790–2799.

24. Bayer, P., A. Arndt, …, J. Becker. 1998. Structure determination of the small ubiquitin-related modifier SUMO-1. *J. Mol. Biol.* 280:275–286.

25. Sippl, M. J., and M. Wiederstein. 2012. Detection of spatial correlations in protein structures and molecular complexes. *Structure.* 20:718–728.

26. Sippl, M. J., and M. Wiederstein. 2008. A note on difficult structure alignment problems. *Bioinformatics.* 24:426–427.

27. Vertegaal, A. C. 2010. SUMO chains: polymeric signals. *Biochem. Soc. Trans.* 38:46–49.

28. Vertegaal, A. C. 2007. Small ubiquitin-related modifiers in chains. *Biochem. Soc. Trans.* 35:1422–1423.

29. Matic, I., M. van Hagen, …, A. C. Vertegaal. 2008. In vivo identification of human small ubiquitin-like modifier polymerization sites by high accuracy mass spectrometry and an in vitro to in vivo strategy. *Mol. Cell. Proteomics.* 7:132–144.

30. Florin, E. L., M. Rief, …, H. E. Gaub. 1995. Sensing specific molecular-interactions with the atomic-force microscope. *Biosens. Bioelectron.* 10:895–901.

31. Bustamante, C., J. F. Marko, …, S. Smith. 1994. Entropic elasticity of lambda-phage DNA. *Science.* 265:1599–1600.

32. Oberhauser, A. F., P. E. Marszalek, …, J. M. Fernandez. 1998. The molecular elasticity of the extracellular matrix protein tenascin. *Nature.* 393:181–185.

33. Rief, M., J. M. Fernandez, and H. E. Gaub. 1998. Elastically coupled two-level systems as a model for biopolymer extensibility. *Phys. Rev. Lett.* 81:4764–4767.

34. Bell, G. I. 1978. Models for the specific adhesion of cells to cells. *Science.* 200:618–627.

35. Evans, E., and K. Ritchie. 1997. Dynamic strength of molecular adhesion bonds. *Biophys. J.* 72:1541–1555.

36. Bustamante, C., Y. R. Chemla, …, D. Izhaky. 2004. Mechanical processes in biochemistry. *Annu. Rev. Biochem.* 73:705–748.

37. Li, M. S., and M. Kouza. 2009. Dependence of protein mechanical unfolding pathways on pulling speeds. *J. Chem. Phys.* 130:145102-1–145102-7.

38. Dietz, H., F. Berkemeier, …, M. Rief. 2006. Anisotropic deformation response of single protein molecules. *Proc. Natl. Acad. Sci. USA.* 103:12724–12728.

39. Bieri, O., J. Wirz, …, T. Kiefhaber. 1999. The speed limit for protein folding measured by triplet-triplet energy transfer. *Proc. Natl. Acad. Sci. USA.* 96:9597–9601.

40. Chyan, C. L., F. C. Lin, …, G. Yang. 2004. Reversible mechanical unfolding of single ubiquitin molecules. *Biophys. J.* 87:3995–4006.

41. Marszalek, P. E., H. Lu, …, J. M. Fernandez. 1999. Mechanical unfolding intermediates in titin modules. *Nature.* 402:100–103.

42. Cao, Y., C. Lam, …, H. Li. 2006. Nonmechanical protein can have significant mechanical stability. *Angew. Chem. Int. Ed. Engl.* 45:642–645.

43. Ainavarapu, S. R., L. Li, …, J. M. Fernandez. 2005. Ligand binding modulates the mechanical stability of dihydrofolate reductase. *Biophys. J.* 89:3337–3344.

44. Best, R. B., B. Li, …, J. Clarke. 2001. Can non-mechanical proteins withstand force? Stretching barnase by atomic force microscopy and molecular dynamics simulation. *Biophys. J.* 81:2344–2356.

45. Sharma, D., G. Feng, …, H. Li. 2008. Stabilization provided by neighboring strands is critical for the mechanical stability of proteins. *Biophys. J.* 95:3935–3942.

46. Li, H., and J. M. Fernandez. 2003. Mechanical design of the first proximal Ig domain of human cardiac titin revealed by single molecule force spectroscopy. *J. Mol. Biol.* 334:75–86.

47. Gao, M., M. Wilmanns, and K. Schulten. 2002. Steered molecular dynamics studies of titin I1 domain unfolding. *Biophys. J.* 83:3435–3445.

48. Mayans, O., J. Wuerges, …, M. Wilmanns. 2001. Structural evidence for a possible role of reversible disulphide bridge formation in the elasticity of the muscle protein titin. *Structure.* 9:331–340.

49. Sadler, D. P., E. Petrik, …, D. J. Brockwell. 2009. Identification of a mechanical rheostat in the hydrophobic core of protein L. *J. Mol. Biol.* 393:237–248.

50. Vehlow, C., H. Stehr, …, M. Lappe. 2011. CMView: interactive contact map visualization and analysis. *Bioinformatics.* 27:1573–1574.

51. Sobolev, V., A. Sorokine, …, M. Edelman. 1999. Automated analysis of interatomic contacts in proteins. *Bioinformatics.* 15:327–332.

52. Schlierf, M., H. Li, and J. M. Fernandez. 2004. The unfolding kinetics of ubiquitin captured with single-molecule force-clamp techniques. *Proc. Natl. Acad. Sci. USA.* 101:7299–7304.

53. Geoffroy, M. C., and R. T. Hay. 2009. An additional role for SUMO in ubiquitin-mediated proteolysis. *Nat. Rev. Mol. Cell Biol.* 10:564–568.

54. van der Veen, A. G., and H. L. Ploegh. 2012. Ubiquitin-like proteins. *Annu. Rev. Biochem.* 81:323–357.

55. Pedrioli, P. G., B. Raught, …, R. Aebersold. 2006. Automated identification of SUMOylation sites using mass spectrometry and SUMmOn pattern recognition software. *Nat. Methods.* 3:533–539.

56. Tatham, M. H., M. C. Geoffroy, …, R. T. Hay. 2008. RNF4 is a poly-SUMO-specific E3 ubiquitin ligase required for arsenic-induced PML degradation. *Nat. Cell Biol.* 10:538–546.

Supporting Material

Single-molecule studies on polySUMO proteins reveal their mechanical flexibility

Hema Chandra Kotamarthi, Riddhi Sharma, and Sri Rama Koti Ainavarapu

Department of Chemical Sciences, Tata Institute of Fundamental Research, Homi Bhabha Road, Colaba, Mumbai 400005, India.
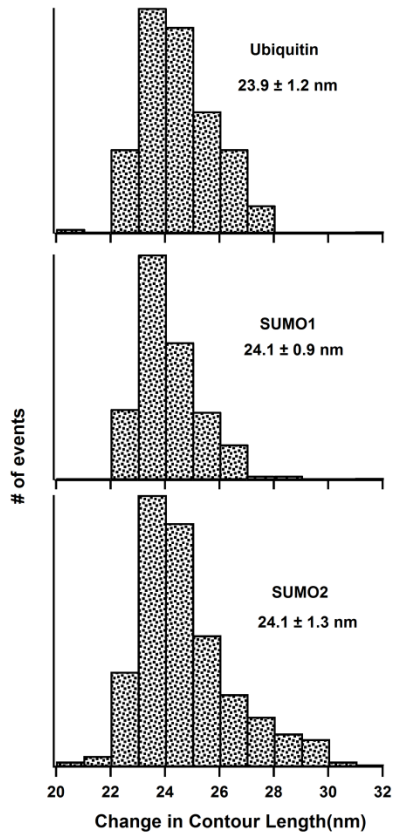
FIGURE S1: Histograms of change in contour length upon unfolding for ubiquitin, SUMO1 and SUMO2. Errors indicate SD.
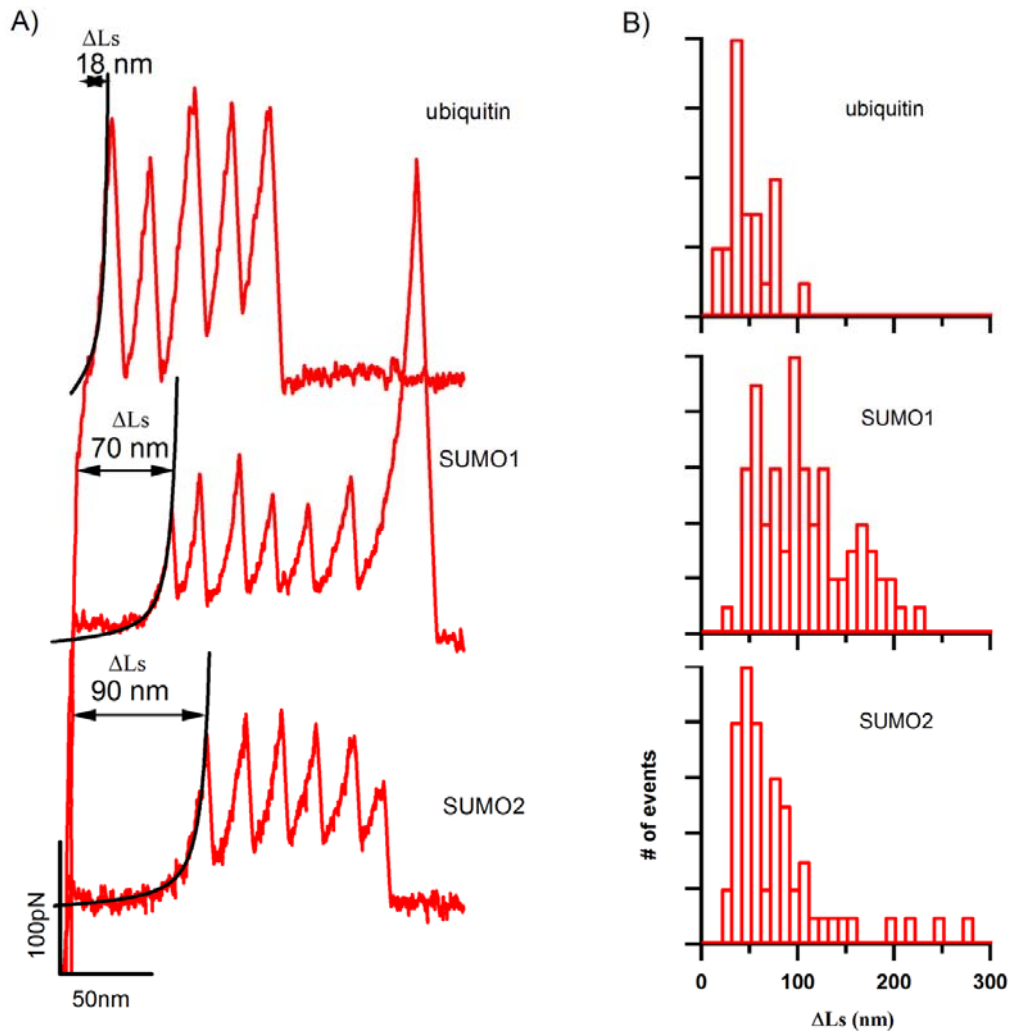
FIGURE S2: A) FX traces of ubiquitin, SUMO1 and SUMO2. Long featureless spacer (ΔLs) prior to the first unfolding force peak was used to estimate the contribution from the structureless region of SUMO1 and SUMO2 proteins. WLC model was used to measure ΔLs as shown here and its histograms are shown in *B*. The spacer lengths are longer for SUMO1 and SUMO2 in comparison to ubiquitin, confirming that the structureless region at the N-terminus of SUMO1 and SUMO2 unravels without any mechanical resistance.
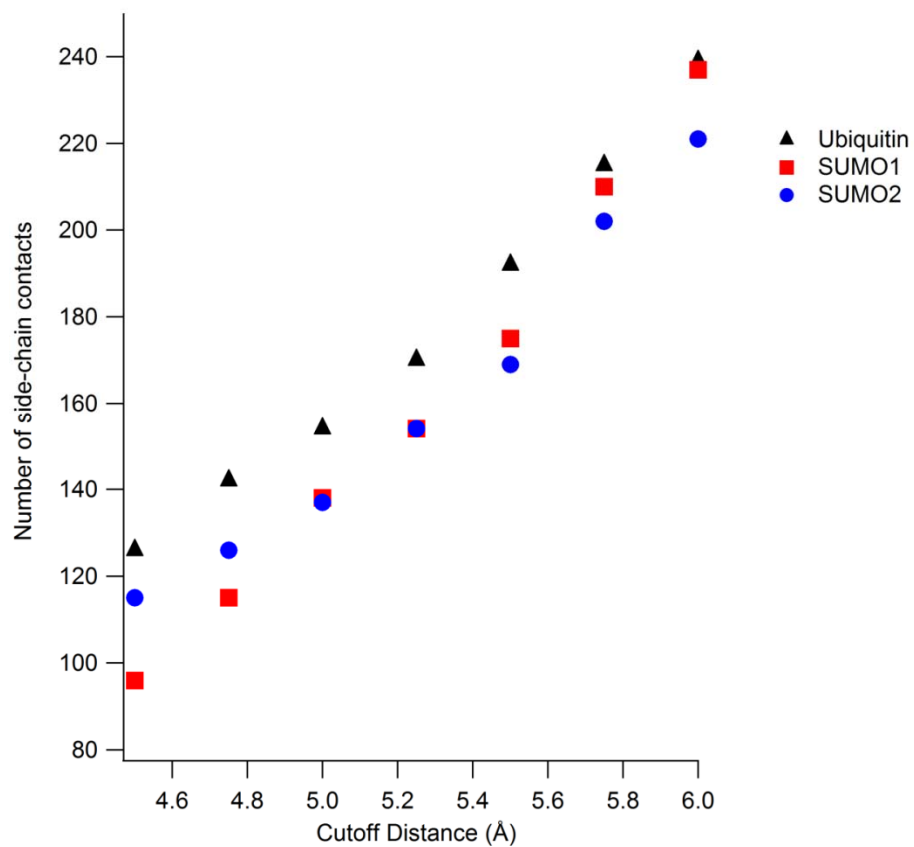
FIGURE S3: Number of side-chain – side-chain contacts versus cut-off distance between residues. The graph shows that the number of contacts in SUMO1 and SUMO2 are always lower than that in ubiquitin.
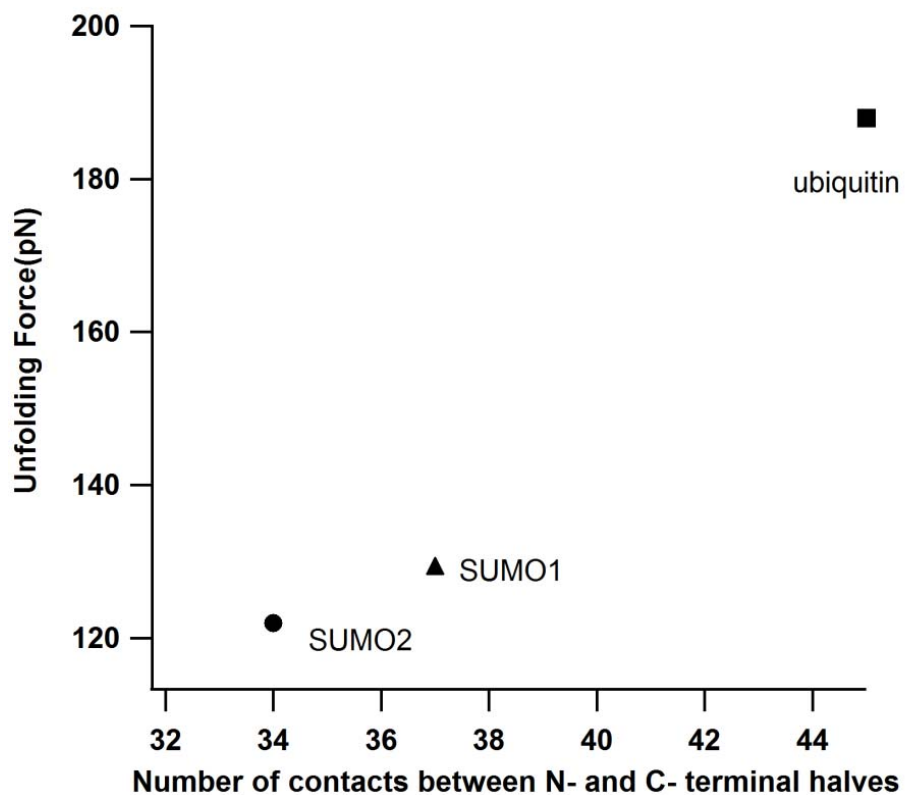
FIGURE S4: Plot of unfolding force versus the number of contacts between the N- and C-terminal halves. Here, the N-terminal (ββα) and C-terminal (ββ) halves of SUMOs and ubiquitin, form the shearing geometry when pulled along N-C direction. The numbers of contacts between these halves also follow the same trend as the overall contacts: ubiquitin (45 contacts) > SUMO1 (37 contacts) > SUMO2 (34 contacts).
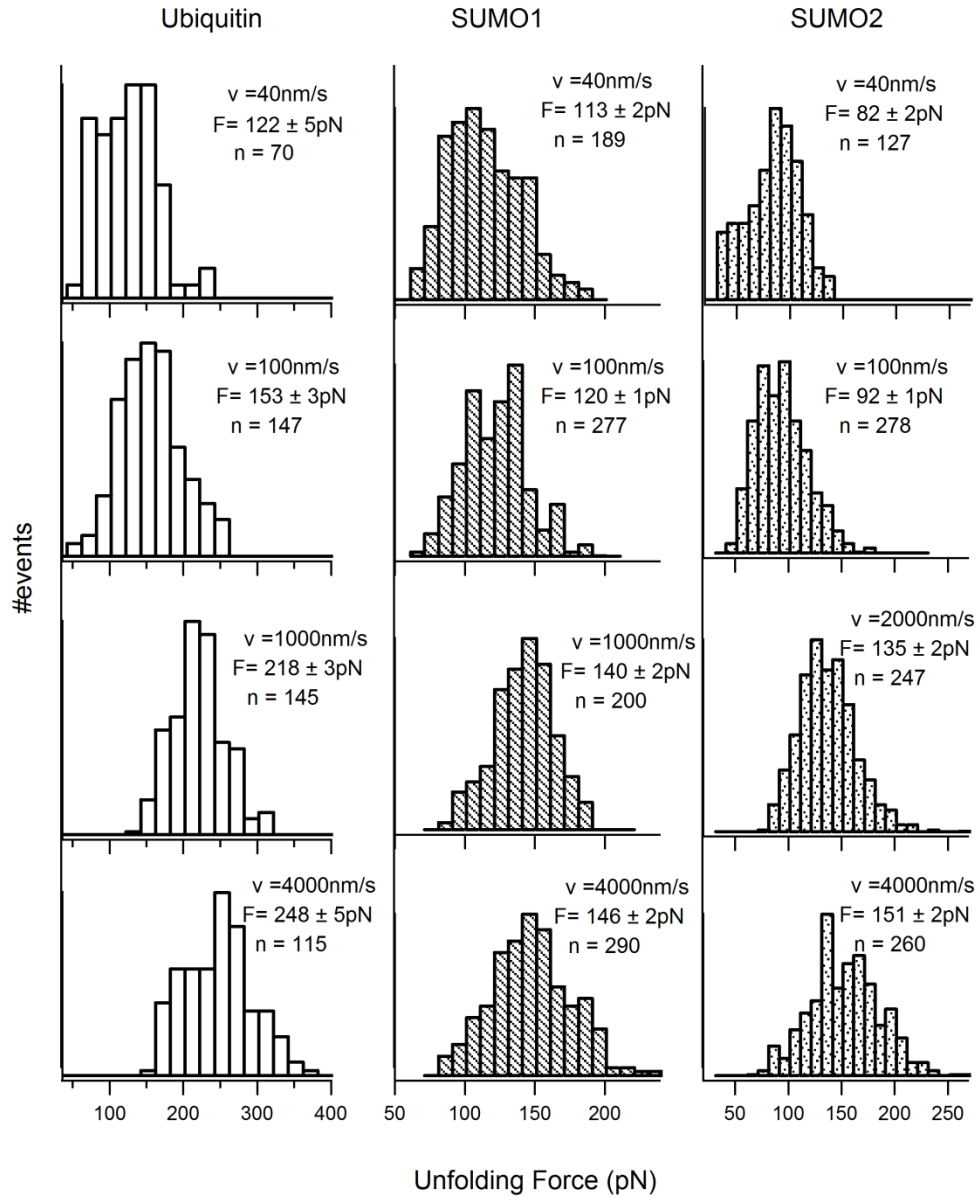
FIGURE S5: Histograms of unfolding forces at different pulling speeds for ubiquitin, SUMO1 and SUMO2. Here, v, F, n represent pulling speed, unfolding force and number of events, respectively. Errors indicate SEM.
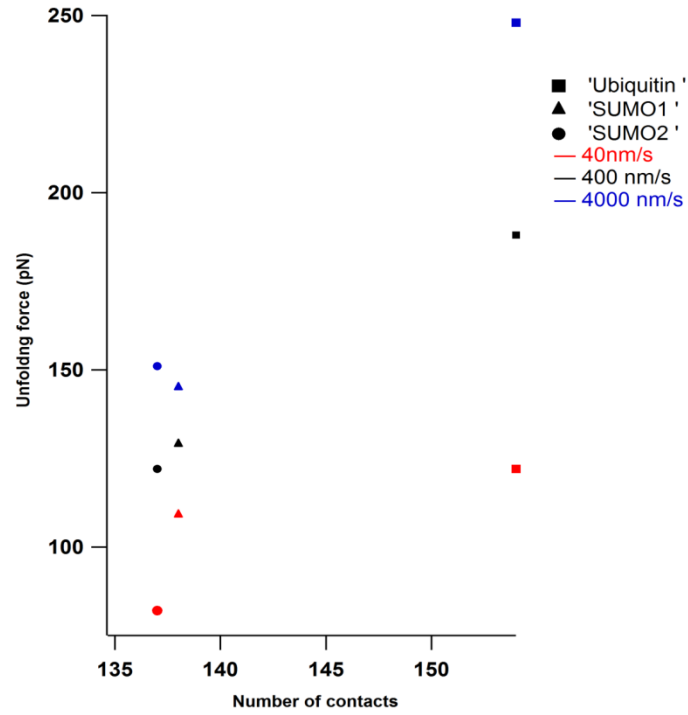
FIGURE S6: Plot of pulling speed dependent unfolding force versus number of contacts. Unfolding forces of SUMO proteins are always lower than that of ubiquitin at all pulling speeds, which correlates with the lower number of contacts for SUMO proteins compared to ubiquitin.

Table S1: Properties of side-chain contacts

	Ubiquitin	SUMO1	SUMO2
Hydrophilic/H-bonds*	59	57	44
Aromatic interactions*	0	2	2
Hydrophobic interactions*	105	100	93
Total Contact Surface Area (Å^2) †	3200	2994	2574

* The side-chain contacts (shown in Fig. 3, Main text) were classified into hydrophilic/H-bonds, aromatic and hydrophobic interactions using CSU software (1). The total number interactions may be higher than that of the total number of contacts (shown in Fig. 3, Main text) for a given protein, as there could be multiple interactions involved in a contact between certain residues of a given protein.
† Contact surface area of non-neighbor residues was also estimated using CSU software (1)

Reference

1.   Sobolev, V., A. Sorokine, J. Prilusky, E. E. Abola, and M. Edelman. 1999. Automated analysis of interatomic contacts in proteins. *Bioinformatics*. 15:327-332.