

Mechanical Unfolding and Refolding of NanoLuc via Single-Molecule Force Spectroscopy and Computer Simulations

Dimitra Apostolidou, Pan Zhang, Weitao Yang,* and Piotr E. Marszalek*



Cite This: *Biomacromolecules* 2022, 23, 5164–5178



Read Online

ACCESS |



Metrics & More

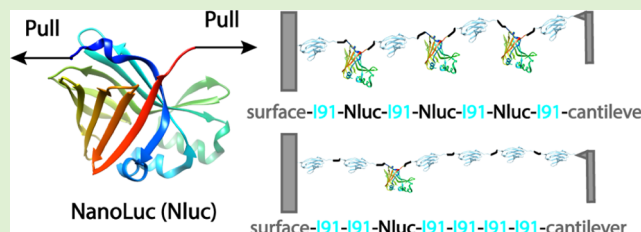


Article Recommendations



Supporting Information

ABSTRACT: A highly bioluminescent protein, NanoLuc (Nluc), has seen numerous applications in biological assays since its creation. We recently engineered a NanoLuc polyprotein that showed high bioluminescence but displayed a strong misfolding propensity after mechanical unfolding. Here, we present our single-molecule force spectroscopy (SMFS) studies by atomic force microscopy (AFM) and steered molecular dynamics (SMD) simulations on two new hybrid protein constructs comprised of Nluc and I91 titin domains, I91-I91-Nluc-I91-I91-I91 (I91₂-Nluc-I91₄) and I91-Nluc-I91-Nluc-I91-Nluc-I91, to characterize the unfolding behavior of Nluc in detail and to further investigate its misfolding properties that we observed earlier for the I91₂-Nluc₃-I91₂ construct. Our SMFS results confirm that Nluc's unfolding proceeds similarly in all constructs; however, Nluc's refolding differs in these constructs, and its misfolding is minimized when Nluc is monomeric or separated by I91 domains. Our simulations on monomeric Nluc, Nluc dyads, and Nluc triads pinpointed the origin of its mechanical stability and captured interesting unfolding intermediates, which we also observed experimentally.



1. INTRODUCTION

NanoLuc (Nluc) is an engineered bioluminescent protein developed by Promega in 2012 with numerous applications in molecular and cellular biology.^{1–18} Although other bioluminescent probes are available, including widely used Firefly Luciferase (Fluc) and Renilla Luciferase (Rluc), Nluc has several important advantages, such as a smaller size, greatly increased brightness, and significantly increased structural and enzymatic stability.^{1,2,6,8,9,12,13,17,18} Nluc originated from a small subunit of the deep sea shrimp *Oplophorus luciferase* (OLuc).^{1,2,19–22} This subunit, OLuc19, is bioluminescent, reacting in an ATP-independent reaction with coelenterazine^{1,2,15,19,20} producing blue light, which is used by the shrimp as a self-defense mechanism against predators. Since OLuc19 was unstable and aggregated when expressed in *Escherichia coli* cells,²² various steps of amino acid mutations and substrate optimization were performed to stabilize this 19 kDa subunit. The result was Nluc and its substrate, furimazine.^{1,2,6,8,9,11–13,15,17–20}

Here, we propose exploiting Nluc for applications beyond its typical biochemical assays by creating new Nluc polyprotein constructs with monomeric Nluc or tandem Nluc repeats in homo- and heteropolyprotein configurations. These new proteins could find many bioengineering applications, e.g., in novel, bioluminescent, force-sensing probes, and elastic (bio)materials with mechanically controlled light output. We also envision exploiting Nluc in more fundamental studies, such as using Nluc as a viable alternative to Fluc in the chaperone-assisted protein refolding as a new substrate.^{23–25} However, in spite of its great potential, Nluc's structure–

mechanics properties have not been adequately characterized so far.

In our first study of Nluc by AFM-based SMFS,⁷ we used a polyprotein Nluc construct composed of three tandem repeats of Nluc flanked by the I91 titin domain resulting in I91-I91-Nluc-Nluc-Nluc-I91-I91 (abbreviated as I91₂-Nluc₃-I91₂). In this construct, the I91 domains serve as pulling handles and provide a single-molecule mechanical fingerprint.^{26–29} This construct is characterized by increased bioluminescence, as compared to monomeric Nluc (by more than 2.5 times), and displays interesting nanomechanical behavior.⁷ When stretched in an AFM, Nluc domains appeared to be moderately strong and unfolded one by one, producing a set of 3 unfolding force peaks of ~70 pN (at stretching speeds of 0.05–0.25 nm/ms) and a contour length increase of ~64 nm, consistent with its 171 amino acid length.⁷ Once mechanically unfolded, Nluc domains surprisingly displayed a strong misfolding propensity in cyclic mechanical stretch and relax measurements, refolding correctly only in ~17% of the cycles.⁷ This “disappointing” mechanical refolding behavior of Nluc is very interesting considering Nluc's small and robust 3D structure^{1,19} and warranted further investigation that could contribute to the

Received: August 10, 2022

Revised: October 27, 2022

Published: November 9, 2022



understanding of proteins' structure–mechanics relationship^{30–35} and could reveal its potential as a chaperone substrate.

We hypothesized that the design of our original construct may have contributed to its misfolding propensity by allowing identical amino acid sequences from different domains to mix while in the unfolded and relaxed state.³⁶ In contrast, other engineered homopolypeptides, such as poly I91 domains of titin or polyubiquitin, generally refold correctly and misfold only occasionally.³⁷ Thus, to examine this hypothesis, we developed new Nluc constructs with multiple Nluc repeats separated from each other by I91 domains and another one with a single Nluc protein flanked by I91 domains and examined these in AFM-based SMFS measurements. These experiments indicate that separating Nluc repeats from each other greatly enhances their ability to refold correctly after mechanical unraveling.

In this report, we also characterize for the first time the mechanical unfolding and refolding properties of Nluc *in silico*. Specifically, we performed coarse-grained unfolding steered molecular dynamics simulations of a monomeric Nluc and on Nluc dyads and triads, as well as multiple all-atom steered molecular dynamics simulations of the solvated Nluc monomer to characterize its mechanical unfolding pathway(s) in detail. We find both experimentally and computationally that in addition to unfolding in a two-state manner, Nluc may also unfold through a mechanical unfolding intermediate. Our observations indicate that Nluc is not only a powerful molecular beacon for *in vivo* cellular studies but also an interesting polypeptide for numerous bioengineering applications and fundamental studies.

2. MATERIALS AND METHODS

2.1. Protein Engineering. Nluc's structure obtained from X-ray crystallography measurements is provided on the PDB website with access number 5IBO. This file was used to acquire the protein's coding sequence from Integrated DNA Technologies Inc (Research Triangle Park, NC). In the PDB file, the first amino acid is methionine, resulting in a total of 171 amino acids. We opted to keep the methionine from the PDB file in our sequence since (a) it is part of the PDB file used in the simulations and (b) we wanted to make sure that methionine was ahead of the protein for generating future constructs. To create the I91₂-Nluc-I91₄ construct, we used a plasmid that was previously developed in our lab,^{24,31} which comprised of two I91 domains, a Firefly Luciferase (Fluc, PDB code 1BA3), and four I91 domains in the order from N- to C-terminus with a double tag at the N-terminus (HisTag-AviTag). We replaced the Fluc sequence with Nluc's sequence in this plasmid; therefore, the final construct was HisTag-AviTag-I91₂-Nluc-I91₄. These changes were performed in the lab using the restriction sites neighboring the protein sequences to be removed. For the HisTag-I91-Nluc-I91-Nluc-I91-Nluc-I91-StrepTag-(II) construct, we used a (pEMI91)²⁶ plasmid that was previously developed in our lab, which carries codon-shuffled nine I91 domains separated by unique cloning sites, unlike the I91 domains in the HisTag-AviTag-I91₂-Nluc-I91₄ construct, which are identical at the codon level. The final construct was synthesized by GenScript Biotech (Piscataway, NJ). The three Nlucs in this construct were also synthesized to have shuffled codons to allow the generation of various primers for sequencing the entire protein gene. These silent mutations were also performed by GenScript Biotech. All constructs mentioned in this study had two cysteines at the C-terminus to enhance the likelihood of the C-terminus binding to the gold-coated sample surface.

2.2. Protein Purification. Both plasmids were transformed with the heat-shock method into BL21-Gold (DE3) competent cells (Agilent, Santa Clara, CA) and plated. Single colonies for both

constructs were grown into starter cultures and then transferred into 500 mL of LB medium to grow further to an OD of 0.8 at 37 °C. Cell cultures were brought to room temperature and induced using 1 mM IPTG for 3 h, followed by the collection of the cell pellet by centrifugation at 4000 rpm for 30 min. The pellet was then resuspended in 5 mL of binding buffer (50 mM NaH₂PO₄/Na₂HPO₄, 300 mM NaCl, 20 mM imidazole, pH 8) and immediately flash frozen in liquid nitrogen and stored at −80 °C until ready for purification. Chemical lysis was usually performed the next day when 25 mL of lysis buffer (50 mM NaH₂PO₄/Na₂HPO₄, 300 mM NaCl, 20 mM imidazole at pH 8.0, protease inhibitor cocktail EDTA-Free, 1 mg/mL lysozyme, 3 units/mL Benzonase, 1 mM DTT, 0.5% Triton X-100) was added to the cells on ice. Both protein constructs were purified by binding to the HisTag column and eluting following the protocol provided by the manufacturer (Ni-NTA Agarose-Qiagen, Germantown, MD). The purified protein was dialyzed using 10k MWCO SnakeSkin Dialysis Tubing (ThermoFisher Scientific, Waltham, MA) in Buffer A (25 mM HEPES, 100 mM KCl, 1 mM DTT) and flash frozen at ~1 mg/mL ready for experiments.

2.3. Atomic Force Microscopy (AFM)-Based Single-Molecule Force Spectroscopy (SMFS). A typical atomic force microscopy (AFM)-based single-molecule force spectroscopy (SMFS) setup is comprised of a sample surface, typically gold coated, and is placed on top of a piezo actuator. The actuator moves the sample surface on the *x*–*y*–*z* axes, with a subnanometer precision on the *z*-axis (pulling direction). Over the surface is a cantilever onto which the protein will nonspecifically adsorb when the cantilever is in contact with the surface. A more detailed description of the setup and a detailed protocol on how we perform AFM-based SMFS experiments have already been reported.³⁸ This protocol was implemented in these studies, which used a custom-built microscope operated by the LabView program (National Instruments, Austin, TX). All spectroscopy experiments were performed using MLCT and SNL cantilevers (Bruker Corporation, Camarillo, CA). The spring constant for these cantilevers was determined at room temperature in Buffer B (25 mM HEPES, 100 mM KCl, 2 mM DTT) using the equipartition theorem³⁹ with the values being ~17 pN/nm (MLCT) and ~80 pN/nm (SNL).³⁹ All experiments were performed at pulling rates of 0.05–0.50 nm/ms, depending on the experiment. Recently gold-evaporated glass slides (Ted Pella, Redding, CA) were used to deposit 50 μL of the diluted protein (0.25–0.5 mg/mL) to incubate for 30 min to allow for nonspecific adsorption of the protein to the gold layer. The protein sample was washed by adding and removing three times 10 μL of buffer B to remove any unbound protein molecules in the solution while the surface was solvated.

2.4. Force–Extension Recordings at Constant Pulling Speed. We operate our AFM force spectrometers in two modes, “down-mode” and “up-mode”. During the “up-mode”, the cantilever is separated from the sample, which is placed on the piezoelectric actuator, and then the piezo controller moves the sample toward the cantilever until they are in contact and then moves back to its initial starting position (no contact) (Figure S1). During the “down-mode”, the cantilever's starting position is in contact with the sample. The piezo controller moves the glass slide with the sample away from the cantilever, separating the two, and then retracts. This separation and retraction can be repeated multiple times to allow cyclic measurements of the protein molecule of interest, enabling its repetitive unfolding and refolding studies. In both cases (up- and down-modes), a molecule is nonspecifically attached to the cantilever during the first contact step. Periodically bringing the cantilever and the sample surface in close proximity during cyclic recording on a single molecule can potentially result in the nonspecific adsorption of another protein molecule(s), which could compromise the experiment. For this reason, we closely monitored the separation distance during such measurements and occasionally compensated for the instrument drift. We note that in these SMFS measurements, the protein is always hydrated and never crosses the water–air interface. Therefore, typically, in our experiments, we used the up-mode to completely extend our constructs (N- to C-terminus), later on labeled as “full-length experiments”, while the down-mode was used for repetitive

Monomeric Nluc: MGSS-HHHHHH-SSG-LVPRGS-HELM-Nluc-TSRT-WSHPQFEK-CCMH
 Dyad Nluc: MGSS-HHHHHH-SSG-LVPRGS-HELM-Nluc-LEGSM-Nluc-ASRT-WSHPQFEK-CCMH
 Triad Nluc: MGSS-HHHHHH-SSG-LVPRGS-HELM-Nluc-LEGSM-Nluc-ASGSM-Nluc-TSRT-WSHPQFEK-CCMH

Figure 1. Sequences for monomeric Nluc, dyad Nluc, and triad Nluc. The six histidines used for His-Tag purification are highlighted in blue, the Thrombin cleavage sites are highlighted in orange, and the Strep-Tag(II) is highlighted in green. All linkers are in black letters. Finally, we identified the methionine from the Nluc PDB file as part of the linkers before Nluc; therefore, we had "...HELM-VFT...ILA-LEGSM-VFT...ILA-ASRT..." for the dyad Nluc.

unfolding and refolding of Nluc, later on labeled as "cyclic experiments." The distance at which the cantilever and the surface needed to be separated during the SMFS measurements was determined by calculating the total length of the fully unfolded protein construct to be studied, using the calibration provided by Dietz and Rief³⁴ with 0.365 nm per amino acid. Nluc is in total 171 amino acids, which would result in $170 \times 0.365 \text{ nm} = 62 \text{ nm}$, and a single I91 titin domain is 89 amino acids, which would result in 32 nm. During the "full-length experiments", our goal was to unfold all protein segments (all Nluc and all I91), so the separation distance was determined to be ~400 to 500 nm, while for the cyclic experiments, our goal was to unfold and refold the Nluc segment, without unfolding I91 domains, which otherwise could potentially interfere with Nluc refolding. This is not difficult to accomplish as Nluc is mechanically weaker than I91 titin domains, so the distance during separation was reduced to ~80 to 100 nm for I91₂-Nluc-I91₄ and ~180 to 230 nm for I91-Nluc-I91-Nluc-I91 constructs.

The force–extension curves were fitted using a worm-like chain (WLC) model with the interpolation formula $\frac{Fp}{k_B T} = \frac{1}{4} \left(1 - \frac{x}{L} \right)^{-2} - \frac{1}{4} + \frac{x}{L}$, where F is the force, p is the persistence length, L is the contour length, x is the extension, and $k_B T = 4.114 \text{ pN nm}$ ⁴⁰ at room temperature, $T = 298 \text{ K}$. Thus, the L and p values of every peak were obtained. The unfolding force, F , and the molecule extension corresponding to the unfolding event, x , were obtained by fitting the raw data and collecting the converging values as the maximum values of the force and length, respectively (see Figure S2 for WLC model fitting on a recording). The statistical values collected from the WLC model are presented below in either histograms, using Sturge's formula for bin size calculations, $k = \lceil \log_2 n \rceil + 1$, where k is the number of bins and n is the sample size,⁴¹ or violin plots.

2.5. Simulations. **2.5.1. Coarse-Grained Simulations of Monomeric, Dyad, and Triad Nluc.** For coarse-grained steered molecular dynamics (SMD) simulations, we used the SMOG server⁴² to generate the C_α native structure-based force field.⁴³ We first performed all-atom simulations for proteins with the CHARMM36 protein force field⁴⁴ and the TIP3P water model.⁴⁵ Minimization was followed by a 1 ns NVT equilibration, with a v-rescale thermostat,⁴⁶ and a 2 ns NPT equilibration, with a Parrinello–Rahman barostat.⁴⁷ The cutoff values for short-range electrostatic interaction and van der Waals interaction were set as 1.2 nm, and long-range electrostatic interaction was calculated by the particle mesh Ewald method.⁴⁸ The integration time step was set as 2 fs. The coarse-grained model was generated from the final equilibrated all-atom model. The best temperature for coarse-grained simulations was determined based on a previous method.²⁴ The protein was denatured with a 300 K coarse-grained simulation, and refolding simulations were performed at 5 different temperatures: 130, 140, 150, 160, and 170 K. We selected 150 K as the best temperature for coarse-grained simulations after plotting the fraction of native contacts since the probabilities for the folded and unfolded states are similar. The sequences of monomeric Nluc, dyad Nluc, and triad Nluc constructs are shown in Figure 1.

We prepared the initial structures based on these sequences. The simulation time was set as 200 ns for each simulation, with 0.5 fs integration time step. The cutoff values for van der Waals interaction was set as 3.0 nm. For the pulling simulation, constant velocity pulling was used by setting the C_α atom of the N-terminal residue as the reference atom and the C_α atom of the C-terminal residue as the pulling atom. The pulling velocity and spring constant were set as 1 nm/ns and 6 pN/nm, respectively. We run 10 simulations for each

system to generate the force–extension plots with average behavior. All simulations were performed in GROMACS 2018.2.⁴⁹

2.5.2. All-Atom Steered Molecular Dynamics (SMD) Simulations of Nluc. For all-atom steered molecular dynamics (SMD) simulations, we used a monomeric Nluc (PDB code 5IBO). In 5IBO, the protein has an additional methionine at the beginning of the N-terminus, which we did not remove, resulting in a total of 171 residues. For all-atom simulations, we did not include linker amino acids that are present in our protein constructs to limit the already big size of the water box needed to accommodate stretched Nluc. We initially used a "large" water box of $7.2 \times 6.2 \times 45.0 \text{ nm}^3$ (system size of ~193 000 atoms) to almost completely unravel Nluc. In later simulations, focused on the initial phase of the forced unfolding, we reduced the size of the water box to $7.7 \times 8.8 \times 15.9 \text{ nm}^3$ (system size of ~103 000 atoms), and $7.7 \times 8.8 \times 9.2 \text{ nm}^3$ (system size of ~59 000 atoms) to perform more simulations to evaluate their variability at the initial phase of stretching at various pulling speeds. In the latter case, the protein was oriented to have N'- and C'-termini aligned on the Z-axis, which we chose as the SMD pulling axis. All simulations were conducted using NAMD 2.12 or 2.14 with CUDA GPU acceleration along with Charm 36 (par_all36m_prot.prm),⁴⁴ the TIP3P water model,⁴⁵ and the ion concentration was set to 150 mM NaCl. Minimization at 0 K was followed by a 5–10 ns equilibration at 298 K prior to pulling simulations. The NPT ensemble was applied for all pulling simulations with a Nöse–Hoover Langevin piston pressure control and Langevin dynamics for temperature control. Periodic boundary conditions were applied. The spring constant was 48.6 pN/nm, and the pulling speed was 0.23–0.25 m/s (system size of ~103 000 atoms) and 0.1 m/s (system size of ~59 000 atoms) for the small water box and 0.125 or 0.25 m/s for the large box (system size of ~193 000 atoms). These velocities are ~10⁶ times greater as compared to the experimental stretching velocity, which is common for all-atom SMD simulations;⁵⁰ however, with recent advances in the graphical processing units and the implementation of supercomputers⁵¹ in simulations, the difference between experiments and simulations has started decreasing to ~10⁴ times greater as compared to experimental stretching.⁵² In all simulations but one, a dummy atom was anchored to the C_α atom of the 171 Ala residue of Nluc, while the C_α atom of methionine (#1) was fixed to perform the constant velocity pulling simulations. In one simulation, we swapped the fixed and SMD atoms to stretch the protein in the opposite direction, with the C_α atom of the Ala (#171) residue being fixed in space and the C_α atom of methionine (#1) being the SMD atom. The water box for this simulation was the smaller water box mentioned previously. In all simulations, the SMD atom was constrained in X and Y directions to mimic the restraint imposed by the AFM cantilever, which does not deform significantly in these directions.

3. RESULTS AND DISCUSSION

Although Nluc has been available and frequently used since 2012,¹ its folding and mechanics have not been studied extensively. From the protein mechanics point of view, the arrangement of β -strands in Nluc is also quite unusual. The first, N-terminal β -strand (strand A, Figure 8) makes a 90° turn to bind to the last C-terminal β -strand (strand J) in an antiparallel geometry. This results in an unusual bond topology, which is neither shear-like^{27,36,53–56} nor zipper-like,^{50,56–59} when the protein is pulled apart by its termini. A somewhat similar topology is displayed by a much smaller src SH3 domain (PDB code 1SRL) that revealed an interesting

mechanical anisotropy in single-molecule force spectroscopy (SMFS) measurements by optical tweezers.⁵⁵ For our initial study of Nluc's mechanics by AFM, we engineered a protein construct with three tandem repeats of Nluc connected through a linker of 5 amino acids, along with titin I91 domains flanking the Nluc proteins resulting in the I91₂-Nluc₃-I91₂.⁷ There are three advantages of making the construct in such an order, with three Nluc modules in the center flanked by two I91 domains on both sides. First, the I91 domains provide pulling handles and some spacing between the Nlucs (protein of interest) and the substrate surface and the AFM cantilever tip, which minimizes nonspecific interactions of the central protein with both surfaces. Second, the I91 domains have been extensively studied mechanically in the past.^{26,60} These domains are known to produce a single unfolding peak (each) with the rupture force, and their contour length increment and persistence length are already well characterized. Therefore, a single full-length recording of the I91₂-Nluc₃-I91₂ construct, which occurs when the construct is stretched from N- to C- terminus, would have four characteristic peaks of I91 domains, demonstrating that the entire construct was stretched and unfolded completely. Since Nluc proteins are in the center of the construct, they too will unfold, producing their own unfolding peaks. As we previously established,⁷ they are mechanically weaker than I91 domains and, therefore, unfold prior to the latter. The third advantage of this construct is that it enables us to record the unfolding of three Nluc proteins from one single molecule, increasing our statistics per molecule and the overall number of unfolding events. Also, registering three similar unfolding events per three Nlucs in a single force spectrogram gives a stronger molecular fingerprint as compared to a single peak, allowing us to establish a "template" to which other measurements can be compared, enhancing data reproducibility. However, due to the proximity of the three Nluc repeats and the probability of misfolding among the three Nlucs during cyclic experiments, we recently developed a new construct that had the same total number of modules but in a different order. The new construct is I91-Nluc-I91-Nluc-I91-Nluc-I91 (Figure S3a). In this case, we again have I91 domains at the N' and C' termini as spacers, preventing Nluc from nonspecifically binding to the surfaces. The other two I91 domains are placed between the three Nlucs acting as additional linkers separating them. We also developed a new construct with a single Nluc to further prevent Nluc's misfolding due to contact with other Nluc modules, a possibility that cannot be excluded even when Nluc modules are separated by I91 domains, because Nluc polypeptide chains can come to close contact during the molecule's relaxation step in unfolding–refolding cyclic experiments. This construct is I91₂-Nluc-I91₄ (Figure S3b). In the end, all three constructs, which were bioluminescent (both constructs with three Nlucs were ~3× brighter than the monomeric Nluc and the single Nluc construct for the same concentrations and experimental conditions), provided an opportunity for a thorough investigation of Nluc's mechanical behavior and the origin of its misfolding propensity.

3.1. Mechanical Unfolding Behavior of Nluc in the Three Constructs I91-Nluc-I91-Nluc-I91-Nluc-I91, I91₂-Nluc-I91₄, and I91₂-Nluc₃-I91₂ Is Similar. We first carried out full-length SMFS stretching experiments on our two new Nluc constructs and compared these results to the force–extension data obtained previously on the I91₂-Nluc₃-I91₂ construct (Figure 2). For the two constructs with three Nluc

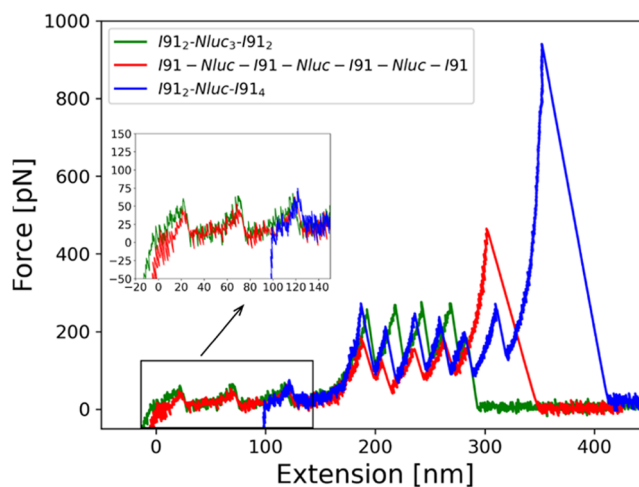


Figure 2. Force–extension curves of the full unfolding of the three constructs: I91₂-Nluc₃-I91₂ (green, pulling speed 0.25 nm/ms),⁷ I91-Nluc-I91-Nluc-I91-Nluc-I91 (red, pulling speed 0.5 nm/ms), and I91₂-Nluc-I91₄ (blue, pulling speed 0.25 nm/ms). The results show that all three constructs have similar behavior when fully stretched. All constructs' unfolding curves begin with the unfolding of the Nluc proteins (highlighted in the black box), followed by the mechanically stronger I91 titin domains. The rupture peaks of both Nluc and I91 for all three constructs demonstrate good overlap, supporting the correct folding of the proteins in all constructs.

modules, we observed three characteristic smaller Nluc peaks preceding the larger I91 peaks. These recordings overlap well, regardless of the specific arrangement of Nluc in the constructs (Figure 2). The construct with a single Nluc module flanked by I91 domains, consistent with its design, displayed a single smaller unfolding force peak. This peak preceded I91 peaks and overlapped well with the Nluc unfolding force peaks of the other two constructs.

These observations suggest that the mechanical stability of Nluc is comparable in all our constructs. These observations also demonstrate that Nluc is in its folded state in all of these constructs and that neither neighboring Nluc modules nor I91 domains significantly affect each other's folding *in vivo*.

We note that the exemplary recordings shown in Figure 2 are representative of measurements performed on full-length proteins that happened to attach to the AFM tip and substrate through their terminal fragments. However, most of our SMFS recordings were "less perfect" in that one or more unfolding peaks from titin I91 domains were missing due to the random nature of the attachment of the protein to the tip (and substrate). Examples of these "partial" recordings are shown in Figure 4b for I91₂-Nluc-I91₄ (dark blue) and Figure S4 for I91-Nluc-I91-Nluc-I91-Nluc-I91 (also see Figure S3 for the schematic examples of end-to-end full-length unfolding and partial unfolding of a polyprotein construct). After establishing the "typical" unfolding behavior of every construct from full-length recordings, partial recordings with fewer I91 domains were also included in our analysis to increase the statistics without undermining the validity of our results (see Figure S17). These partial recordings overlap well with the full-length traces. However, they may slightly differ from each other (e.g., the magnitude of the unfolding force or small variations in the contour length increments). This is due to the stochastic nature of the forced unfolding, measurements' random errors, and some more systematic errors resulting from, e.g., the variation in the pulling angle and AFM cantilever's spring

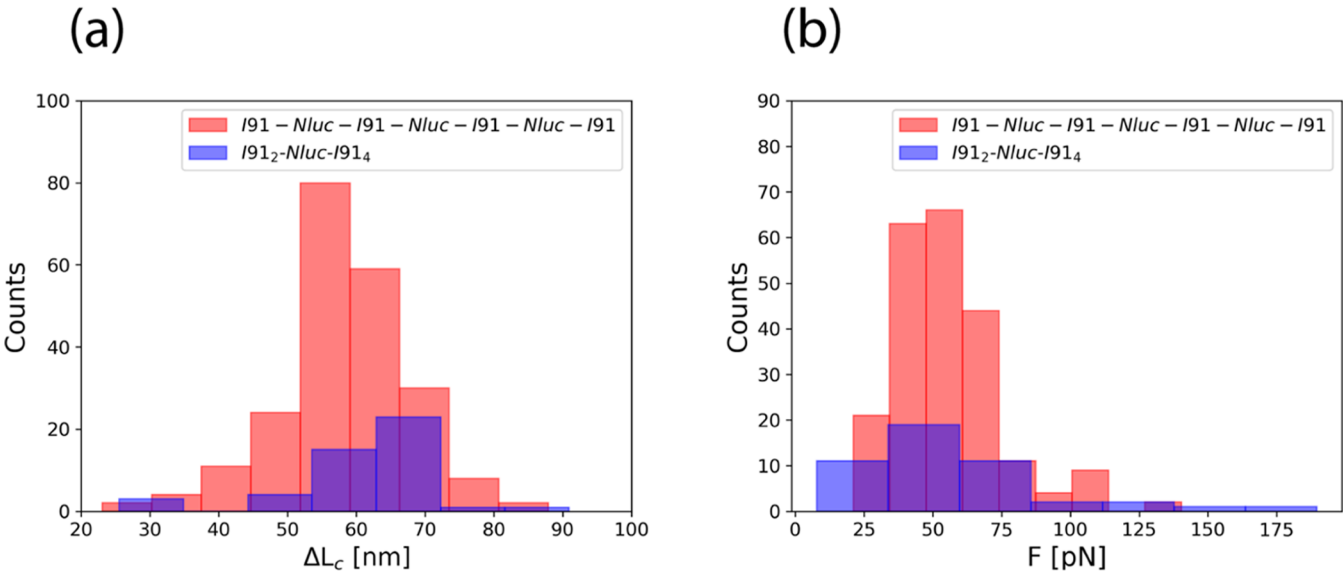


Figure 3. Histograms of (a) the contour length increment (ΔL_c) and (b) the unfolding forces (F) of Nluc after the WLC fitting of the recordings collected by full-length SMFS stretching experiments of I91-Nluc-I91-Nluc-I91-Nluc-I91 (red) and I91₂-Nluc-I91₄ (blue).

Table 1. Unfolding Forces (F), Contour Length Increment (ΔL_c), Persistence Length (p), and Number of Recordings (n) for the Three Constructs^a

construct	F [pN]	ΔL_c [nm]	p [nm]	n
I91 ₂ -Nluc ₃ -I91 ₂	72 ± 3	64 ± 1	0.4	54 ^b
I91-Nluc-I91-Nluc-I91-Nluc-I91	55 ± 1	58 ± 1	0.50 ± 0.02	220
I91 ₂ -Nluc-I91 ₄	58 ± 5	61 ± 2	0.57 ± 0.08	47

^aEvery value has the mean and standard error of the mean. ^bReference 7.

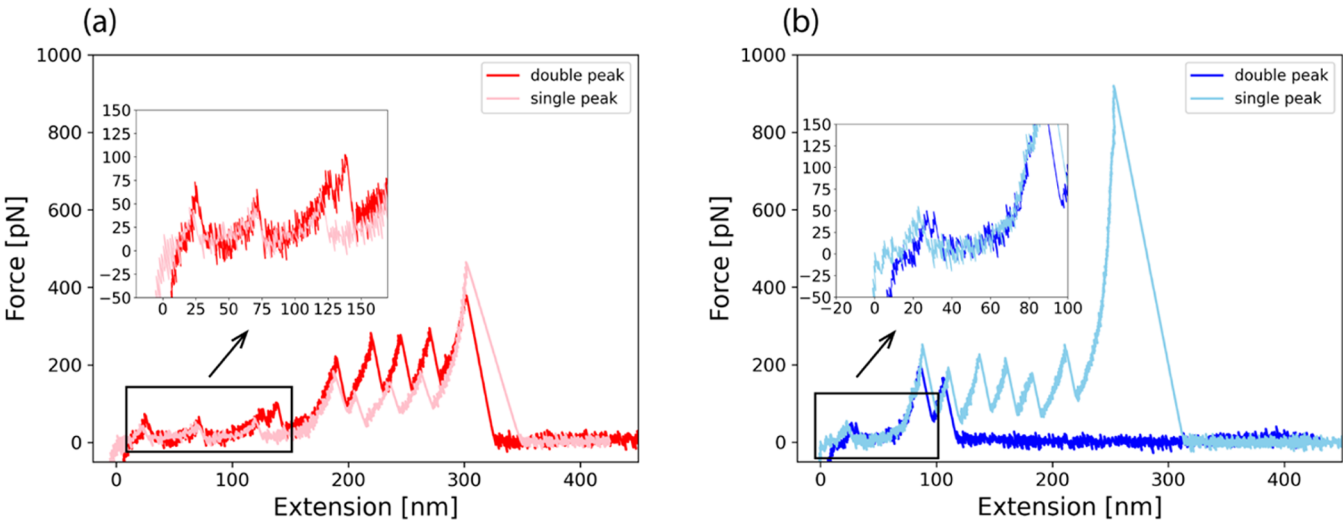


Figure 4. Recordings of full unfolding with double peaks for a Nluc protein for (a) I91-Nluc-I91-Nluc-I91-Nluc-I91 and (b) I91₂-Nluc-I91₄. For I91-Nluc-I91-Nluc-I91-Nluc-I91, only one of the Nlucs shows double peaks (red), while the other two are single peaks (pulling speed 0.25 nm/ms). For I91₂-Nluc-I91₄, the Nluc (blue) shows two peaks (pulling speed 0.3 nm/ms). Both constructs were overlapped with recordings of the same construct in which all Nluc peaks were single (with pink (a) and light blue (b)). For I91-Nluc-I91-Nluc-I91-Nluc-I91 colored in pink, the pulling speed was 0.5 nm/ms, and for I91₂-Nluc-I91₄ colored in light blue, the pulling speed was 0.25 nm/ms.

constant calibration uncertainty.^{61,62} Recordings showing large nonspecific adhesive interactions, represented by irregular force peaks that usually appear at the beginning of the extension process, as well as recordings suggestive of more than one molecule being stretched (e.g., as indicated by significant irregularities in I91 unfolding force peaks) were not included in the analysis.

Overall, our results indicate that Nluc unfolds with similar mechanical characteristics in all constructs studied. Histograms of contour length increments (nm) and unfolding forces (pN) of Nluc are presented in Figure 3 for I91-Nluc-I91-Nluc-I91-Nluc-I91 and I91₂-Nluc-I91₄. As we can see from our histograms, the distributions are in good agreement.

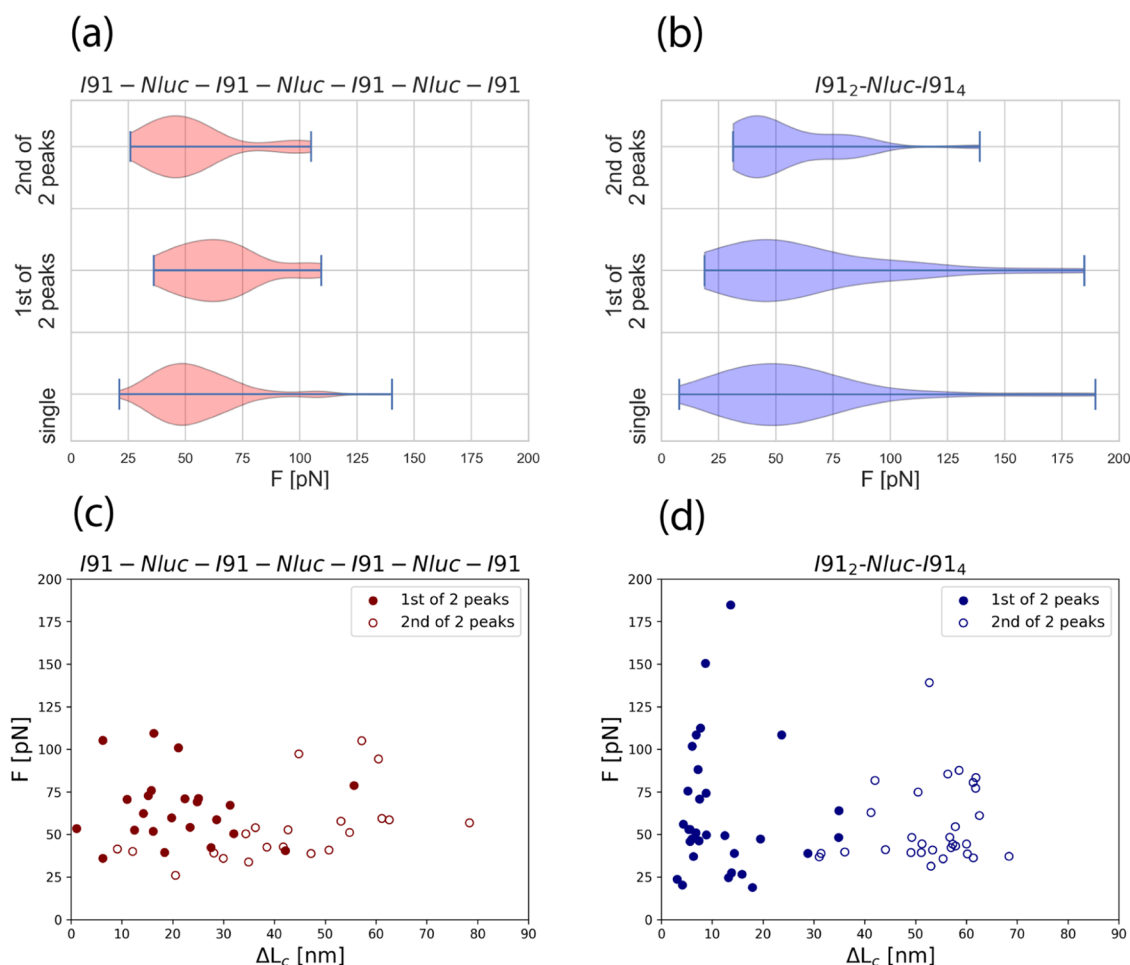


Figure 5. Violin plots of unfolding force, F , for single peaks, 1st of 2 peaks of double peaks, and 2nd of 2 peaks of double peaks for (a) I91-Nluc-I91-Nluc-I91-Nluc-I91 and (b) I91₂-Nluc-I91₄. Scatter plots of unfolding force, F , versus contour length increment, ΔL_c , showing the distribution and clustering of the 1st and 2nd peaks for double peaks for (c) I91-Nluc-I91-Nluc-I91-Nluc-I91 and (d) I91₂-Nluc-I91₄. For I91-Nluc-I91-Nluc-I91 ($n = 32$) and for I91₂-Nluc-I91₄ ($n = 30$).

Calculated mean values and standard errors of the mean of the unfolding force, contour length increment, and persistence length are shown in Table 1. We also compared our results to those of the I91₂-Nluc₃-I91₂ construct.⁷ We observe a general agreement in values among the three constructs establishing that Nluc is able to fold correctly in all three constructs. Somewhat surprisingly, the unfolding forces for the two new constructs are on average lower than for I91₂-Nluc₃-I91₂ (statistically significant when compared individually to both constructs in this study, see Supporting Information Section 9). However, the contour length increment and persistence length seem to be more consistent among the constructs. The numbers of recordings, n , are shown for each construct in Table 1.⁷

3.2. Nluc Unfolding Shows Single and Double Force Peaks. In our initial study of I91₂-Nluc₃-I91₂, we have reported that Nluc modules unfold with a single force peak, each in an all-or-none fashion. However, in this study on I91-Nluc-I91-Nluc-I91-Nluc-I91 and I91₂-Nluc-I91₄, we also occasionally observed two closely spaced peaks, whereas in other recordings, typically only one peak was present. This observation suggests that a stepwise unfolding of Nluc is also possible. Overlapping recordings with single and double peaks (Figure 4) suggest that the contour length increment following a single unfolding event is similar to the overall length

increment resulting from double-peak events. This behavior was observed for both constructs, as shown in Figure 4. A full-length recording for I91₂-Nluc-I91₄ with double peaks and all six I91 titin domains unfolded can be found in Figure S5.

When analyzing all Nluc unfolding force peaks in our SMFS recordings, we determined that 91% ($n = 220$) are single and 9% ($n = 21$) are double peaks for I91-Nluc-I91-Nluc-I91-Nluc-I91, while 61% ($n = 47$) are single and 39% ($n = 30$) are double peaks for I91₂-Nluc-I91₄. The reason behind this difference is currently not clear and warrants further studies.

In Figure 5a,b, we present an overlay of violin plots of the rupture (unfolding) force for single peaks and 1st and 2nd peaks of double peaks for I91-Nluc-I91-Nluc-I91-Nluc-I91 and I91₂-Nluc-I91₄, respectively. We conclude that all of these events cover the same range of forces. Figure 5c,d shows scatter plots of the unfolding forces against the corresponding contour length increment for both 1st and 2nd peaks of one Nluc molecule. These demonstrate that the clustering of the rupture forces is within the same range for I91-Nluc-I91-Nluc-I91-Nluc-I91 and I91₂-Nluc-I91₄, respectively.

We performed a separate analysis for both 1st and 2nd peaks of double-peak events to calculate the mean and standard error of the mean of F , ΔL_c , and p for both constructs, and the results are shown in Table 2. Both constructs demonstrate significant agreement in the results of unfolding forces and

Table 2. Unfolding Forces (F), Contour Length Increment (ΔL_c), Persistence Length (p) for 1st and 2nd of Double Peaks per Nluc Protein from Full-Unfolding Recordings of I91-Nluc-I91-Nluc-I91-Nluc-I91 and I91₂-Nluc-I91₄^a

construct	I91-Nluc-I91-Nluc-I91-Nluc-I91	I91 ₂ -Nluc-I91 ₄
F [pN] 1st of 2 peaks	65 ± 4	63 ± 7
F [pN] 2nd of 2 peaks	53 ± 5	55 ± 4
ΔL_c [nm] 1st of 2 peaks	21 ± 3	12 ± 2
ΔL_c [nm] 2nd of 2 peaks	43 ± 4	53 ± 2
p [nm] 1st of 2 peaks	0.51 ± 0.05	0.50 ± 0.07
p [nm] 2nd of 2 peaks	0.55 ± 0.09	0.47 ± 0.07

^aEvery value has the mean and standard error of the mean.

persistence length, with some interesting differences in the contour length increment. In I91-Nluc-I91-Nluc-I91-Nluc-I91, the 1st peak is associated with $\Delta L_c = 21 \pm 3$ nm, while for I91₂-Nluc-I91₄, it is $\Delta L_c = 12 \pm 2$ nm. To further understand the elements contributing to the generation of single and double peaks for a single Nluc protein, we performed SMD simulations, and our findings are discussed in Section 3.4.

Finally, we confirmed that both peaks within double peaks likely originate from the same Nluc module. This was done by comparing the sum of the ΔL_c of the double peaks (1st and 2nd) to the ΔL_c of the single peaks for both constructs. In Figure 6, histograms of the two distributions are presented. As expected from the results in Table 2, the combined contour length increment of the double peaks matches with one of the single peaks. As expected, this is further supporting the fact that the two split peaks and the single peak reflect the unfolding of one Nluc protein.

3.3. Cyclic Unfolding/Refolding SMFS Experiments of Nluc Demonstrate That It Can Refold Robustly When Separated by I91 Domains.

Cyclic experiments were also performed, during which the protein molecules were stretched and relaxed in a repetitive manner using down-mode operation. These experiments are consequential in providing significant information on the ability of proteins to refold correctly or their tendency to misfold. Therefore, these are valuable in collecting information on unfolding/refolding pathways of the protein and its possible intermediates. Our

previous study on I91₂-Nluc₃-I91₂ showed that in 56.2% of refolding cycles, Nluc was not refolding correctly (no peaks); in 16.9%, correct refolding was observed; and in 26.9%, misfolding occurred.⁷

For the I91-Nluc-I91-Nluc-I91-Nluc-I91 construct, we registered $n = 118$ cyclic recordings with 61.9% ($n = 73$) having one to three Nluc native or near-native peaks, 16.9% ($n = 20$) being misfolded cases, and 21.2% ($n = 25$) being unfolded cases (no peaks). We further split the misfolded cases into two subcategories, with 11.0% of total recordings ($n = 13$) having misfolding that did not show any Nluc peaks and 5.9% ($n = 7$) of total recordings that had one or two Nluc peaks that were native or near-native with some misfolding for the other Nluc(s) modules. Similarly, for I91₂-Nluc-I91₄, we split our $n = 93$ cyclic recordings into three categories. The first one had Nluc peaks being native or near-native (40.9%, $n = 38$), the second had misfolding with no clear Nluc peaks (24.7%, $n = 23$), and the third had no refolding of Nluc (34.4%, $n = 32$). For the I91₂-Nluc-I91₄ construct, we also registered $n = 21$ cyclic recordings that indicated significant instrumental drift. For those $n = 21$ recordings, it was impossible to reliably determine the extension of Nluc and its unfolding/refolding status. Thus, those recordings were not included in the above analysis. Some drift was also observed in SMFS measurements of the I91-Nluc-I91-Nluc-I91-Nluc-I91 construct. However, since there were three Nluc peaks at the beginning of the recording, the limited drift still allowed us most of the time to observe one or two Nluc unfolding behavior if they occurred. This is a good demonstration of the advantage of having multiple repeats of the same protein to be studied in cyclic measurements. In Figure 7, we demonstrate examples of cyclic measurements for both constructs.

Our findings are very interesting given our previous studies of the I91₂-Nluc₃-I91₂ construct showed that 16.9% of recordings refolded correctly, while 56.2% would not refold and 26.9% would misfold. In both of our new constructs, the percentage of correctly refolding Nluc modules was significantly higher as compared to the percentage of misfolding modules. Given the unfolded states were observed in the majority of the recordings for the I91₂-Nluc₃-I91₂ construct, it supports our hypothesis that when Nlucs are linked to each

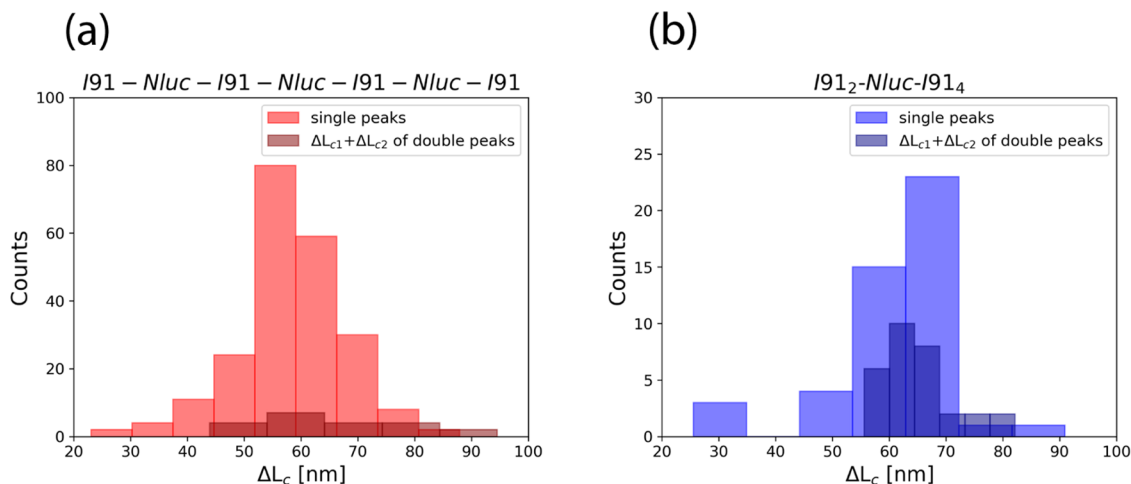


Figure 6. Histograms of the combined contour length increments of the double peaks compared to those of the single peaks for (a) I91-Nluc-I91-Nluc-I91-Nluc-I91 and (b) I91₂-Nluc-I91₄. The distributions show that $\Delta L_{c, \text{single}}$ is overlapping well with $\Delta L_{c1} + \Delta L_{c2}$, indicating that the two peaks of the double peaks originate from the unfolding of one Nluc protein.

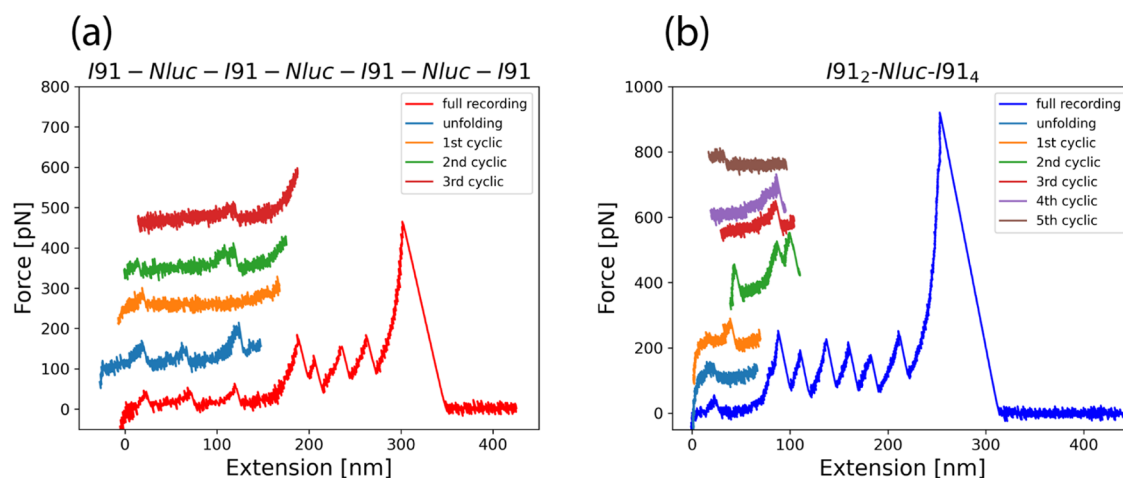


Figure 7. Force–extension recordings of cyclic measurements for (a) I91-Nluc-I91-Nluc-I91-Nluc-I91 and (b) I91₂-Nluc-I91₄. For both (a) and (b), we are comparing the partial cyclic measurements with a representative of the full unfolding for both constructs. All unfolding and cyclic measurements were for one single molecule that was continually unfolded, relaxed, and refolded until it detached from either the cantilever or the surface. For (a) and (b), we color-coded the 1st partial unfolding curve with light blue, followed by the various cycles in order. We do not demonstrate the relaxation curves since they did not show any forces being generated. For I91-Nluc-I91-Nluc-I91-Nluc-I91 in (a), the full recording had a pulling speed of 0.5 nm/ms, the unfolding curve had 0.25 nm/ms (blue), the 1st cyclic measurement had 0.25 nm/ms (orange), the 2nd cyclic measurement had 0.25 nm/ms (green), and the 3rd cyclic measurement had 0.15 nm/ms (red). For I91₂-Nluc-I91₄ in (b), the full recording had a pulling speed of 0.25 nm/ms, the unfolding curve had 0.1 nm/ms (blue), the 1st cyclic measurement had 0.1 nm/ms (orange), the 2nd cyclic measurement had 0.1 nm/ms (green), the 3rd cyclic measurement had 0.1 nm/ms (red), the 4th cyclic measurement had 0.1 nm/ms (purple), and the 5th cyclic measurement had 0.1 nm/ms (brown).

other, they are prone to misfolding in cyclic unfolding/refolding measurements and that the possibility is largely eliminated when Nluc(s) is flanked by other proteins. This hypothesis was further supported by our preliminary *in silico* studies of thermal unfolding and refolding of three tandem Nluc domains, which suggest that the possibility of domain swapping between neighboring Nlucs may be involved in this misfolding process (see Supporting Information Section 12). In contrast, single Nluc is able to refold correctly in our simulations. These studies are a work in progress, and more investigation is required.

Our analysis of the mean and standard error of the mean of the unfolding forces, contour length increment, and persistence length for the cyclic measurements are shown in Table 3. Again, as observed in the full-length recordings of the two constructs, we recorded single and double peaks per Nluc molecule during these measurements. For I91₂-Nluc-I91₄,

Table 3. Unfolding Forces (F), Contour Length Increment (ΔL_c), and Persistence Length (p) for Single and Double Peaks (Separated as 1st and 2nd of 2 Peaks) per Nluc Protein from Cyclic Unfolding Recordings of I91-Nluc-I91-Nluc-I91-Nluc-I91 and I91₂-Nluc-I91₄^a

construct	I91-Nluc-I91-Nluc-I91-Nluc-I91	I91 ₂ -Nluc-I91 ₄
F [pN] single peaks	50 ± 1	46 ± 3
F [pN] 1st of 2 peaks	57 ± 6	48 ± 5
F [pN] 2nd of 2 peaks	49 ± 5	34 ± 6
ΔL_c [nm] single peaks	59 ± 2	54 ± 3
ΔL_c [nm] 1st of 2 peaks	17 ± 3	18 ± 2
ΔL_c [nm] 2nd of 2 peaks	42 ± 6	42 ± 4
p [nm] single peaks	0.48 ± 0.03	0.34 ± 0.06
p [nm] 1st of 2 peaks	0.41 ± 0.07	0.33 ± 0.04
p [nm] 2nd of 2 peaks	0.51 ± 0.09	0.18 ± 0.03

^aEvery value has the mean and standard error of the mean.

single peaks were 55% ($n = 22$) and double peaks were 40% ($n = 16$), while for I91-Nluc-I91-Nluc-I91-Nluc-I91, 91% ($n = 117$) represent single peaks and 9% ($n = 12$) reveal double peaks. Interestingly, these percentages for both constructs are similar to the ones we found in noncyclic, full-unfolding experiments. Also, the results in Table 3 show similar behavior between the two constructs, regardless of single or double peaks per Nluc protein. Interestingly, the ΔL_c of the first peak for both constructs is 17–18 nm, which is an in-between value of the two constructs in full-length recordings ($\Delta L_c = 21$ nm and $\Delta L_c = 12$ nm). Additionally, we observe that during cyclic measurements, the forces and persistence length give slightly smaller values for both constructs.

3.4. Mechanical Unfolding of Nluc *In Silico*. To gain an insight into structural rearrangements of Nluc under force, we performed numerous coarse-grained (CG) and all-atom steered molecular dynamics (SMD) simulations. Here, we start by introducing Nluc's structure as obtained from the PDB file SIBO and proceed by describing the results from the CG-SMD calculations, followed by all-atom SMD simulations.

Nluc is a very interesting protein with an unusual sequence having little homology to other bioluminescent proteins.⁶³ Its distinct 3D structure (PDB code SIBO, Figure 8b) bears a strong resemblance to that of intracellular lipid binding proteins (iLBPs),^{1,63} whose fold is highly conserved despite their low (20%) sequence similarity.⁶⁴ Nluc, like iLBPs, has antiparallel “ β -clams” like β -strand barrel structures that are arranged in a nearly orthogonal orientation. Also, like iLBPs, Nluc has a large solvent-filled cavity in the center of its barrel. However, unlike iLBPs, Nluc has three additional helices (marked in Figure 8a as a0, aIII, and aIV consistent with the iLBPs nomenclature) along with an additional small β -sheet (marked D' in Figure 8a). Additionally, overlapping the structure of Nluc and that of a known iLBP (see Figure S6), it is evident that in Nluc, the aII helix is displaced so that it is now parallel to the aIII helix rather than the aI helix as seen in

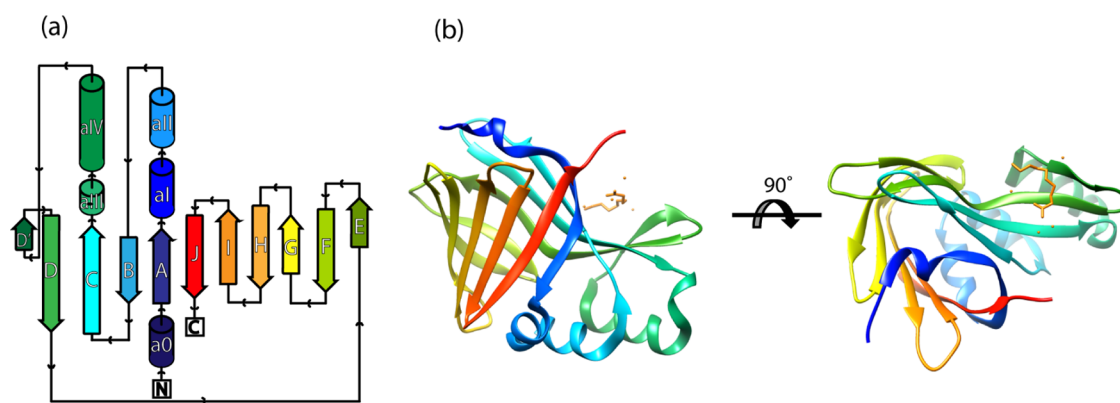


Figure 8. (a) 2D topology of the Nluc protein (adapted from PBDSum) and (b) the 3D structure of Nluc from X-ray experiments (PDB code SIBO) in rainbow depiction (N-terminus is colored blue and C-terminus is colored red). We also did a 90° rotation of the protein to show the N- to C-terminus distance. The gold sticks and balls show the decanoic acid and water molecules nearby, which were part of the PDB file and were present during the X-ray crystallography experiments. The decanoic acid is used here as a model of where furimazine would be interacting with Nluc.

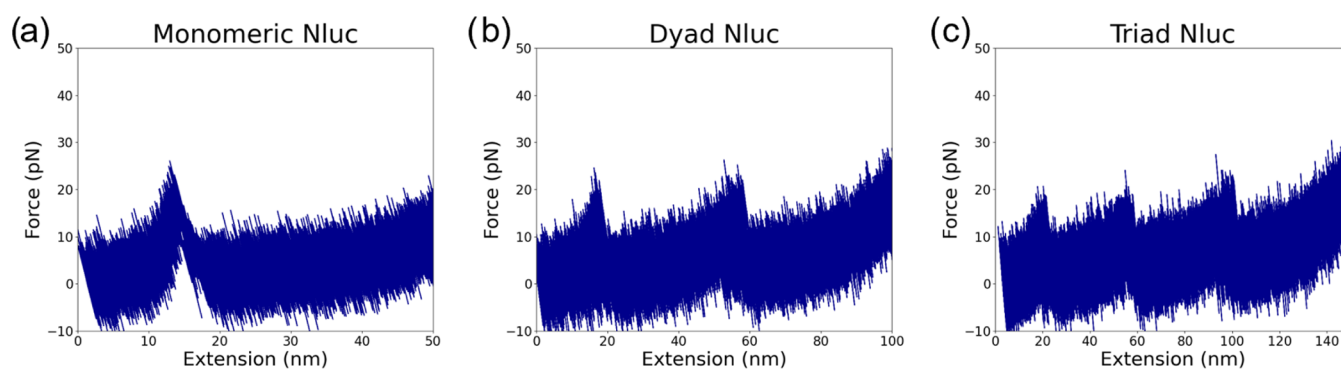


Figure 9. Force–extension plots from coarse-grained simulations of (a) monomeric Nluc, (b) dyad Nluc, and (c) triad Nluc.

iLBPs, which is probably to accommodate for the extra β -sheet and helices found in Nluc.⁶³ Additionally, unlike iLBPs, Nluc is continuously hydrogen bonded with the β D and β E sheets forming hydrogen bonds (see Figure S7). Most importantly, in Nluc, like iLBPs, the β A strand has a unique twist allowing it to form H-bonds with both β B and β J (as seen in Figure 8 colored in dark blue), which seems to stabilize this protein structure.⁶⁴

3.4.1. Coarse-Grained SMD Simulations of Nluc. For the coarse-grained SMD simulations of monomeric Nluc, dyad Nluc, and triad Nluc, the force–extension plots are shown in Figure 9. Similar to SMFS data, the force–extension plots of monomeric Nluc and triad Nluc have 1 major peak and 3 major peaks, respectively. Also, Nluc dyad (not studied by SMFS), consistent with its design, shows 2 force peaks. The peak of monomeric Nluc appears at an extension of around 15 nm, the peaks of dyad Nluc appear at extensions of around 15 and 55 nm, and the peaks of triad Nluc appear at extensions of around 15, 55, and 95 nm, respectively. After dissecting the process of coarse-grained simulations, every peak corresponds to the unfolding of a Nluc domain.

Next, we examined the unfolding process of monomeric Nluc in more detail. The initial CG structure of monomeric Nluc is shown in Figure 10a, which is the folded state. The secondary structures are colored red for the α -helix and blue for the β -sheet. In the first 15 ns, only loops are stretched, and the secondary structures remain well folded. At the simulation time of 15.9 ns, the β -strand J is detached from the β -sheet

near the C-terminus (see Figure 10b), which is the first unfolding event of the secondary structure. At the simulation time of 16.6 ns, the β -strand A is unfolded, as shown in Figure 10c. During the next 100 ps, the β -strands B, C, D, E, F, G, H, and I and almost all secondary structures are unfolded, resulting in the structure displayed in Figure 10d. From Figure 10b–10d, we could observe that the main unfolding process is quite fast and is completed in 800 ps. The whole unfolding process corresponds to the single large peak in the force–extension plot. After 16.7 ns, the whole protein unfolds steadily. As shown in Figure 10e, all secondary structure elements are unfolded, and the protein structure becomes more and more aligned with the stretching direction at longer simulation times.

From the coarse-grained simulations of monomeric Nluc, dyad Nluc, and triad Nluc, we conclude that the coarse-grained simulations provide force–extension plots that are generally consistent with our SMFS data. More details of the unfolding process can be gained by analyzing all-atom SMD simulations.⁶⁵

3.4.2. All-Atom SMD Simulations of Nluc. Our experimental results have shown that Nluc mostly unfolds, producing one rupture force peak, but in a significant number of experiments, particularly when using the monomeric Nluc construct, we also observed two rupture force peaks per Nluc. The possibility of two different unfolding pathways has not been observed in CG-SMD simulations, which captured a single major force peak for each Nluc. This result may reflect

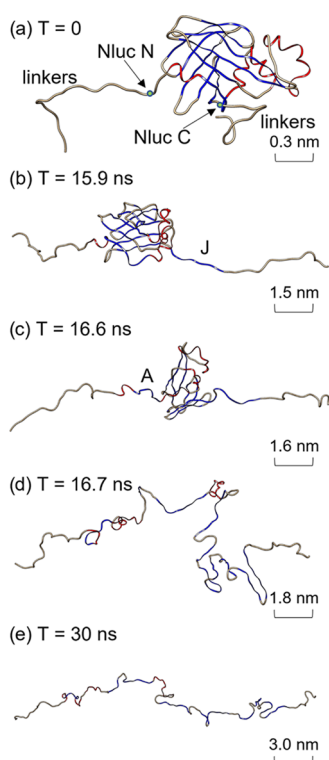


Figure 10. Series of snapshots along the CG-SMD unfolding trajectory at different times. The red color represents the α -helix secondary structures, and the blue color represents the β -strand secondary structures.

the limitation of the CG approach that averages out details of intramolecular interactions. To acquire atomistic-level information about Nluc's mechanical unfolding pathways, we performed all-atom, explicit solvent SMD simulations. These may provide greater insight into the response of structure-stabilizing interactions during the mechanical unfolding of proteins.

3.4.2.1. "Large Box" SMD Simulations. We performed in total six all-atom, explicit solvent SMD simulations of Nluc in a large water box containing $\sim 193\,000$ atoms (see "Nluc large box simulation.avi" for one of the simulations). Its longest dimension was ~ 440 Å to accommodate the unfolding of most of the Nluc chain, with its expected contour length of 620 Å. Each simulation took ~ 2 months on a simple Linux desktop computer equipped with, e.g., an I9 8-core processor and a GTX 980Ti or GTX 1070 GPU. An overlay of six force–extension curves determined from those simulations is shown in Figure 11. In contrast to CG-SMD simulations, all-atom simulations show richer force spectrograms with a greater number of force peaks of varying magnitude, occurring at slightly different extensions in different SMD runs. In Figure 11, we grouped these major force peaks, which stand out from the simulation noise, as A, B, and C according to the extensions at which they occur of ~ 4 , 11, and 32 nm, respectively. As in AFM-based SMFS experiments, the molecular extension is defined here as the projection of the end-to-end vector on the pulling axis.

We examined the molecular trajectories of these SMD simulations and focused on the time points at which major force peaks occurred for every simulation. Consistent in all but one simulation (that did not produce a strong first peak A but produced a prominent peak C), the maximum force of the

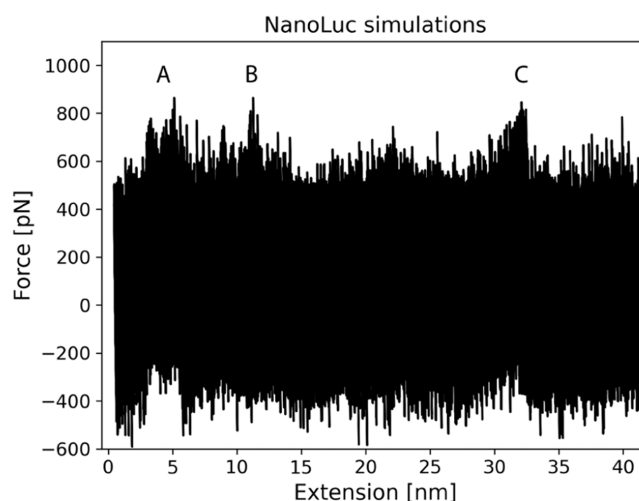


Figure 11. Force–extension profiles for the six all-atom SMD simulations with constant velocity in the large box. With A, B, and C, we identify force peaks generated by the breaking of the secondary structures, located ~ 4 , 11, and 32 nm, respectively.

force peaks in group A occurs when the molecule is highly tensed just before the strand βJ starts peeling off from its β -sheet. Then, the force decreases when the hydrogen bonds of the βJ with the βA and βI strands break (Figure S12). This observation is consistent with the main result from our GC simulations.

Up to this point, both the CG and all-atom simulations suggest that twisting of the βA strand, allowing it to make contact with the two β -sheets, provides rather moderate mechanical stability to Nluc when stretched by its N' and C' termini. This strength reflects the structural arrangement of Nluc, positioning only a small part of its bonds in a stretch-resisting geometry against the applied force (α 0 helix, βA – βJ interactions). At the same time, the βJ – βI bonds are oriented in a "weak" zipper-like geometry and break one by one under force. The appearance of the first major unfolding force peak (A) early in the stretching process is consistent with our experimental SMFS measurements.

Following the first unzipping and detachment of the βJ strand, Nluc's unfolding follows somewhat different pathways in different SMD runs. However, common events in the second phase of the unfolding process involve strand βA sliding out of the structure (consistent with our GC results) while breaking its stabilizing interactions with strands βB and βI . These events occur at Nluc's extension around 10 nm, thus likely contributing to the group of force peaks marked as B (see Figure S13). What happens next is the separation of the N-terminal part (lobe) of the protein from the C-terminal part, following the rupture of the H-bonds between strands βD and βE . βD remains with the N-terminal part, and βE connected to βF moves together with the β sheet still formed by strands G, H, and I. This event occurs at extensions between major peaks B and C in SMD force spectrograms but by itself does not generate a significant force peak. Further events vary between SMD runs, suggesting a highly stochastic nature of this forced unfolding, with the majority of the trajectories indicating the unfolding of the N-terminal part preceding the unfolding of the remaining sheet formed by strands βE – βI . However, one of the simulations that was run to complete the unfolding of the protein showed the βE – βI sheet breaking first, with the βB – βD breaking later, along with the unfolding of the helices

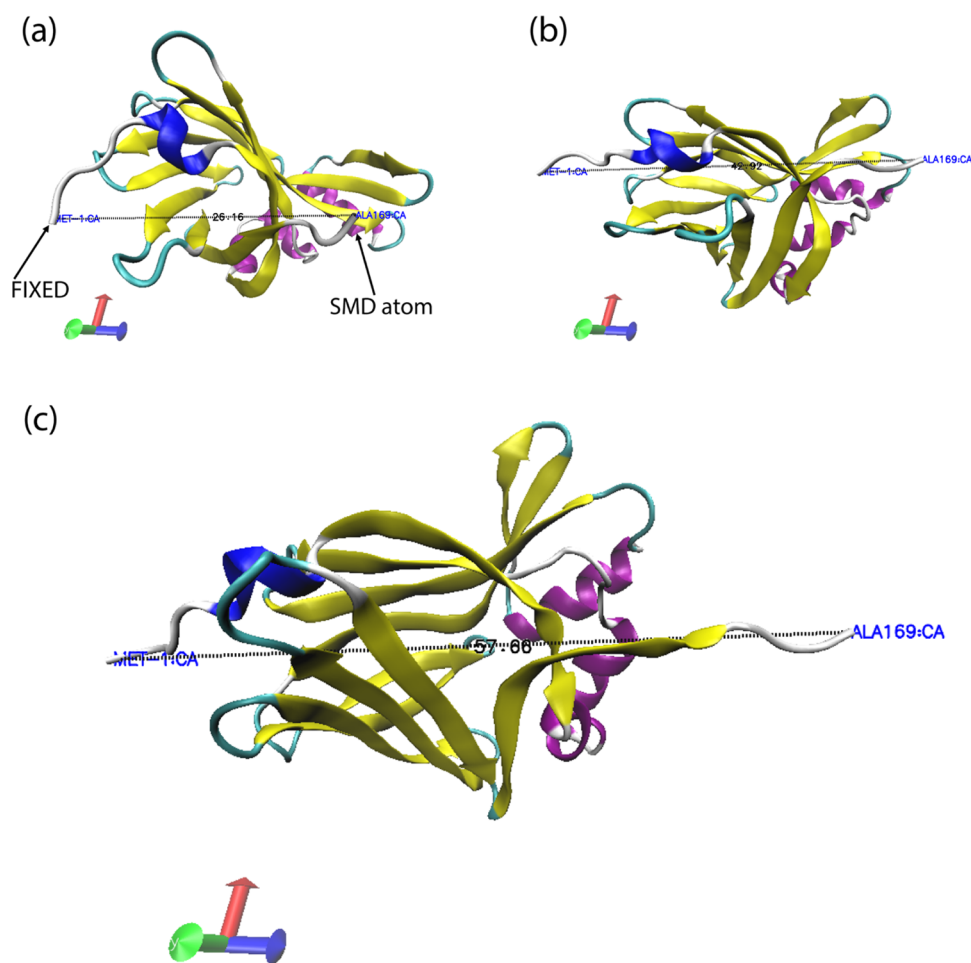


Figure 12. Movie frames for smaller box simulations. (a) Starting position of the simulation at time zero, (b) frame before peak A is generated (42.92 Å), and (c) frame after peak A (57.66 Å). With peak A we refer to the peak in Figure S8. In this simulation, the starting distance was 2.616 nm (26.16 Å). We labeled the residues in blue and the end-to-end distance in black dashed lines.

between these sheets. Thus, force peak C may be related to the unfolding of either N- or C-terminal sections of Nluc (Figure S14) and more large box complete unfolding simulations are required to determine the prevalence of either pathway. It is interesting to note that all-atom SMD simulations identified three major force peaks when unfolding a single Nluc protein. AFM-based SMFS measurements identified only two force peaks, the first appearing early in the stretching process and the second during the later phase of the unfolding process. Of note is the fact that the second force peak, recorded for our construct with a single Nluc domain, tends to appear closer to the first peak (ΔL_c between both peaks is ~ 12 nm; see Table 2). In comparison, the second unfolding peak of the construct with three Nluc domains tends to be separated from the first peak by ΔL_c of 21 nm. Thus, it is possible that the prevalent non-two-state unfolding pathway proceeds as $N \rightarrow A \rightarrow B \rightarrow U$ for a single Nluc domain, while for the triple Nluc construct, it is $N \rightarrow A \rightarrow C \rightarrow U$, where N is the native state, A represents the first major unfolding event, B and C represent unfolding intermediates, and U represents the unfolded state. However, more SMFS and SMD experiments/simulations are needed to verify this conjecture. We note that changing the pulling direction from $N \rightarrow C$ to the opposite ($C \rightarrow N$) did not significantly change the unfolding pathway (Figures S10 and S11), with the detachment of strand βJ being the major force-induced event.

3.4.2.2. “Small Box” SMD Simulations. To further characterize early events during Nluc stretching, we carried out 10 additional all-atom SMD simulations at a higher sampling rate (trajectories saved every 200–250 steps, instead of 1000 steps in large box simulations), albeit in a smaller water box of ~ 103 and 59 kiloatoms (see “Nluc small box simulation.avi” for one of the 103 kiloatom simulations). The reduction of the water box was primarily motivated by an attempt to keep the trajectory file under 10 GB per simulation. For the 103 kiloatom simulations, in five out of seven SMD simulations, the major unfolding event contributing to the first force peak around 4.5–6.5 nm is the detachment of the βJ strand (see Figure 12), as in our large box simulations. However, in two other simulations, the first force-induced event was a significant stretching of the N-terminal α -helix ($\alpha 0$ in Figure 8a), which contributed to the maximum tension at the first force peak. This was followed by a “delayed” detachment of the βJ strand, which started at extensions greater than 5.5 nm (see Figure S8 for force–extension results). For the 59 kiloatom simulations, in all three simulations, the main event appears to be the detachment of the βJ strand from the βA and βI strands (see Figure S16 for movie frames of the simulation). The unfolding peak was observed around 1.5 nm (Figure S9). Another interesting detail that emerged from the small box simulations was the formation of transient stabilizing interactions between a stretch

of the β A strand comprising residues Thr13 to Asn17 with a stretch of the β I strand comprising residues Leu149 to Arg152 (see Figure S15a). These interactions followed the complete detachment of the β J strand and the realignment of the structure against the force. This event occurred in four out of seven simulations. Therefore, breaking this transient structure contributed to the second force peak occurring at the extension of ~ 12 nm (see Figure S15). Thus, the formation of a new transient mini β -sheet involving β A and β I may be the origin of the first unfolding intermediate, also observed in our large box simulations. Please note that the initial orientation of Nluc in the large and small box simulations was slightly different. In the small box simulations, the molecule was fully aligned in the direction of the force at the onset. Thus, the molecular extensions in both sets of simulations are not numerically identical.

4. CONCLUSIONS

All in all, we have studied the Nluc protein in two different constructs and compared our results to the previously reported experiments. All constructs appeared to have similar unfolding behavior with similar unfolding forces and contour length increment values. Also, our findings confirmed our hypothesis that Nluc misfolds when linked to itself. This misfolding behavior of Nluc was concluded when comparing cyclic unfolding and refolding experiments of the I91-Nluc-I91-Nluc-I91-Nluc-I91 and I91₂-Nluc-I91₄ constructs to the I91₂-Nluc₃-I91₂ construct. Our findings from the former two constructs showed that most of the time Nluc was able to successfully refold with similar percentages for both constructs, supporting a similar trend between the two constructs. This finding is in contrast to the ones from the I91₂-Nluc₃-I91₂ construct, which mainly showed misfolding of Nluc. Therefore, we conclude that while all three constructs show similarity in the unfolding pathway, the refolding pathway differs due to the separation or not of the Nluc proteins. We are currently working on identifying the origin of this misfolding behavior of Nluc in more detail, studying its *in silico* refolding behavior after thermal denaturation. Our preliminary results suggest that domain swapping in Nanoluc triads may be contributing to Nluc's misfolding. Additionally, we demonstrated for the first time that Nluc can unfold, producing a single peak or double peaks per molecule. This feature is more present in the I91₂-Nluc-I91₄ construct, with 39% of the recordings demonstrating double peaks over the 9% in I91-Nluc-I91-Nluc-I91-Nluc-I91. Additionally, we performed SMD simulations and provided for the first time valuable insight into the pathways in which Nluc unfolds. The main unfolding event of Nluc was highly supported from all simulations to be the breaking of the hydrogen bonds of the β J sheet with the β A and β I sheets. This is consistent with previous studies, which have shown how deletion in iLBPs of the last three amino acids of the β J sheet resulted in the opening of the barrel. This caused the proteins to have a less compact non-native state with lower thermodynamic stability and higher susceptibility to proteolytic attack.⁶⁴ We also propose, based on our simulations, that the unfolding of the β B– β D sheets or breaking of the transient mini β -sheet formed by the β A and β I strands could be the most likely cause of the 2nd unfolding peak of Nluc. Therefore, all of our simulations provided great insight into the possible unfolding pathways of the Nluc protein, which for a variety of unfolding speeds (0.1–0.25 m/s) and direction of pulling (N- or C-terminus) showed similar behavior. We are currently

investigating in more detail Nluc's possible misfolding pathways, which could further elucidate the misfolding behavior as we demonstrated in our cyclic experiments.

Therefore, our experimental results provide the foundation to explore Nluc in potential bioluminescent, force-sensing, and elastic (bio)materials by simultaneously utilizing its mechanical and bioluminescent properties. We are reporting an interesting behavior of both Nluc's unfolding and refolding/misfolding, which can determine the use of various protein construct designs based on their applications. An example of these applications could be a simultaneous detection of Nluc's luminescence during SMFS measurements, a feat yet to be achieved at a single-molecule level. Very difficult simultaneous SMFS and fluorescence measurements have already been reported.^{33,66} Last but not least, we envision exploiting Nluc as a viable alternative to Fluc in protein refolding and chaperone mechanistic studies.^{23–25}

■ ASSOCIATED CONTENT

SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.biomac.2c00997>.

Schematic of up-mode and down-mode operation of the AFM-based SMFS setup; WLC model fitting of recordings and ΔL_c calculations; visual representation of pulling of various constructs during AFM-based SMFS experiments; partial recordings of I91-Nluc-I91-Nluc-I91-Nluc-I91; example of double peaks of I91₂-Nluc-I91₄ with all I91 domains present; iLBPs and Nluc structural comparison; AFM-based SMFS statistical analysis of the I91 titin domains from the full-length recordings of I91-Nluc-I91-Nluc-I91-Nluc-I91 and I91₂-Nluc-I91₄; small box force–extension profiles; small box force–extension profiles at 0.1 nm/ms; small box opposite pulling direction force–extension profiles; movie frames for small box with opposite pulling direction; movie frames for a large box; movie frames for a small box; movie frames for a small box at 0.25 times slower pulling speed; statistical analysis; selection process of SMFS recordings; overlapping of recordings; *in silico* thermal denaturation of single and triad Nluc (PDF)

Nluc large box simulation (AVI)

Nluc small box simulation (AVI)

■ AUTHOR INFORMATION

Corresponding Authors

Weitao Yang – Department of Chemistry, Duke University, Durham, North Carolina 27708, United States; orcid.org/0000-0001-5576-2828; Email: weitao.yang@duke.edu

Piotr E. Marszalek – Department of Mechanical Engineering and Materials Science, Duke University, Durham, North Carolina 27708, United States; orcid.org/0000-0001-7948-8034; Email: pemar@duke.edu

Authors

Dimitra Apostolidou – Department of Mechanical Engineering and Materials Science, Duke University, Durham, North Carolina 27708, United States; orcid.org/0000-0003-3001-4019

Pan Zhang — Department of Chemistry, Duke University,
Durham, North Carolina 27708, United States

Complete contact information is available at:

<https://pubs.acs.org/10.1021/acs.biomac.2c00997>

Author Contributions

This manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation (Grant Number MCB 1817556).

ABBREVIATIONS

Nluc, NanoLuc; Fluc, firefly luciferase; Rluc, renilla luciferase; Oluc, oplophorus luciferase; AFM, atomic force microscopy; SMFS, single-molecule force spectroscopy; SMD, steered molecular dynamics; PDB, Protein Data Bank; CG, coarse-grained; iLBPs, intracellular lipid binding proteins

REFERENCES

- (1) Hall, M. P.; Unch, J.; Binkowski, B. F.; Valley, M. P.; Butler, B. L.; Wood, M. G.; Otto, P.; Zimmerman, K.; Vidugiris, G.; Machleidt, T.; Robers, M. B.; Benink, H. A.; Eggers, C. T.; Slater, M. R.; Meisenheimer, P. L.; Klaubert, D. H.; Fan, F.; Encell, L. P.; Wood, K. V. Engineered Luciferase Reporter from a Deep Sea Shrimp Utilizing a Novel Imidazopyrazinone Substrate. *ACS Chem. Biol.* **2012**, *7*, 1848–1857.
- (2) Dale, N. C.; Johnstone, E. K. M. M.; White, C. W.; Pfleger, K. D. G. G. NanoBRET: The Bright Future of Proximity-Based Assays. *Front. Bioeng. Biotechnol.* **2019**, *7*, 56.
- (3) Lackner, D. H.; Carré, A.; Guzzardo, P. M.; Banning, C.; Mangena, R.; Henley, T.; Oberndorfer, S.; Gapp, B. V.; Nijman, S. M. B.; Brummelkamp, T. R.; Bürckstümmer, T. A Generic Strategy for CRISPR-Cas9-Mediated Gene Tagging. *Nat. Commun.* **2015**, *6*, No. 10237.
- (4) Lebbink, R. J.; de Jong, D. C. M.; Wolters, F.; Kruse, E. M.; van Ham, P. M.; Wiertz, E. J. H. J.; Nijhuis, M. A Combinational CRISPR/Cas9 Gene-Editing Approach Can Halt HIV Replication and Prevent Viral Escape. *Sci. Rep.* **2017**, *7*, No. 41968.
- (5) Rozbeh, R.; Forchhammer, K. Split NanoLuc Technology Allows Quantitation of Interactions between PII Protein and Its Receptors with Unprecedented Sensitivity and Reveals Transient Interactions. *Sci. Rep.* **2021**, *11*, No. 12535.
- (6) Dixon, A. S.; Schwinn, M. K.; Hall, M. P.; Zimmerman, K.; Otto, P.; Lubben, T. H.; Butler, B. L.; Binkowski, B. F.; Machleidt, T.; Kirkland, T. A.; Wood, M. G.; Eggers, C. T.; Encell, L. P.; Wood, K. V. NanoLuc Complementation Reporter Optimized for Accurate Measurement of Protein Interactions in Cells. *ACS Chem. Biol.* **2016**, *11*, 400–408.
- (7) Ding, Y.; Apostolidou, D.; Marszalek, P. Mechanical Stability of a Small, Highly-Luminescent Engineered Protein NanoLuc. *Int. J. Mol. Sci.* **2021**, *22*, 55.
- (8) Cevenini, L.; Calabretta, M. M.; Lopreside, A.; Tarantino, G.; Tassoni, A.; Ferri, M.; Roda, A.; Michelini, E. Exploiting NanoLuc Luciferase for Smartphone-Based Bioluminescence Cell Biosensor for (Anti)-Inflammatory Activity and Toxicity. *Anal. Bioanal. Chem.* **2016**, *408*, 8859–8868.
- (9) Loh, J. M. S.; Proft, T. Comparison of Firefly Luciferase and NanoLuc Luciferase for Biophotonic Labeling of Group A *Streptococcus*. *Biotechnol. Lett.* **2014**, *36*, 829–834.
- (10) Salameh, J. W.; Kumar, S.; Rivera-Cruz, C. M.; Figueiredo, M. L. A Second-Generation Nanoluc-IL27 Fusion Cytokine for Targeted-Gene-Therapy Applications. *Bioengineering* **2022**, *9*, 77.
- (11) Machleidt, T.; Woodroffe, C. C.; Schwinn, M. K.; Méndez, J.; Robers, M. B.; Zimmerman, K.; Otto, P.; Daniels, D. L.; Kirkland, T. A.; Wood, K. V. NanoBRET—A Novel BRET Platform for the Analysis of Protein–Protein Interactions. *ACS Chem. Biol.* **2015**, *10*, 1797–1804.
- (12) Kamkaew, A.; Sun, H.; England, C. G.; Cheng, L.; Liu, Z.; Cai, W. Quantum Dot-NanoLuc Bioluminescence Resonance Energy Transfer Enables Tumor Imaging and Lymph Node Mapping: In Vivo. *Chem. Commun.* **2016**, *52*, 6997–7000.
- (13) Goyet, E.; Bouquier, N.; Ollendorff, V.; Perroy, J. Fast and High Resolution Single-Cell BRET Imaging. *Sci. Rep.* **2016**, *6*, No. 28231.
- (14) Robers, M. B.; Dart, M. L.; Woodroffe, C. C.; Zimprich, C. A.; Kirkland, T. A.; Machleidt, T.; Kupcho, K. R.; Levin, S.; Hartnett, J. R.; Zimmerman, K.; Niles, A. L.; Ohana, R. F.; Daniels, D. L.; Slater, M.; Wood, M. G.; Cong, M.; Cheng, Y.-Q.; Wood, K. V. Target Engagement and Drug Residence Time Can Be Observed in Living Cells with BRET. *Nat. Commun.* **2015**, *6*, No. 10091.
- (15) Germain-Genevois, C.; Garandeau, O.; Couillaud, F. Detection of Brain Tumors and Systemic Metastases Using NanoLuc and Fluc for Dual Reporter Imaging. *Mol. Imaging Biol.* **2016**, *18*, 62–69.
- (16) Kincaid, V. A.; Wang, H.; Sondgeroth, C. A.; Torio, E. A.; Ressler, V. T.; Fitzgerald, C.; Hall, M. P.; Hurst, R.; Wood, M. G.; Gilden, J. K.; Kirkland, T. A.; Lazar, D.; Chia-Chang, H.; Encell, L. P.; Machleidt, T.; Zhou, W.; Dart, M. L. Simple, Rapid Chemical Labeling and Screening of Antibodies with Luminescent Peptides. *ACS Chem. Biol.* **2022**, *17*, 2179–2187.
- (17) Gaspar, N.; Walker, J. R.; Zambito, G.; Marella-Panth, K.; Lowik, C.; Kirkland, T. A.; Mezzanotte, L. Evaluation of NanoLuc Substrates for Bioluminescence Imaging of Transferred Cells in Mice. *J. Photochem. Photobiol. B* **2021**, *216*, No. 112128.
- (18) Su, Y.; Walker, J. R.; Park, Y.; Smith, T. P.; Liu, L. X.; Hall, M. P.; Labanieh, L.; Hurst, R.; Wang, D. C.; Encell, L. P.; Kim, N.; Zhang, F.; Kay, M. A.; Casey, K. M.; Majzner, R. G.; Cochran, J. R.; Mackall, C. L.; Kirkland, T. A.; Lin, M. Z. Novel NanoLuc Substrates Enable Bright Two-Population Bioluminescence Imaging in Animals. *Nat. Methods* **2020**, *17*, 852–860.
- (19) England, C. G.; Ehlerding, E. B.; Cai, W. NanoLuc: A Small Luciferase Is Brightening Up the Field of Bioluminescence. *Bioconjugate Chem.* **2016**, *27*, 1175–1187.
- (20) Yeh, H. W.; Xiong, Y.; Wu, T.; Chen, M.; Ji, A.; Li, X.; Ai, H. W. ATP-Independent Bioluminescent Reporter Variants to Improve in Vivo Imaging. *ACS Chem. Biol.* **2019**, *14*, 959–965.
- (21) Inouye, S.; Watanabe, K.; Nakamura, H.; Shimomura, O. Secretional Luciferase of the Luminous Shrimp *Oplophorus gracilirostris*: cDNA Cloning of a Novel Imidazopyrazinone Luciferase. *FEBS Lett.* **2000**, *481*, 19–25.
- (22) Inouye, S.; Sasaki, S. Overexpression, Purification and Characterization of the Catalytic Component of *Oplophorus gracilirostris* Luciferase in the Deep-Sea Shrimp, *Oplophorus gracilirostris*. *Protein Expr. Purif.* **2007**, *56*, 261–268.
- (23) Schröder, H.; Langer, T.; Hartl, F. U.; Bukau, B. DnaK, DnaJ and GrpE Form a Cellular Chaperone Machinery Capable of Repairing Heat-Induced Protein Damage. *EMBO J.* **1993**, *12*, 4137–4144.
- (24) Scholl, Z. N.; Yang, W.; Marszalek, P. E. Chaperones Rescue Luciferase Folding by Separating Its Domains. *J. Biol. Chem.* **2014**, *289*, 28607–28618.
- (25) Mayer, M. P.; Gierasch, L. M. Recent Advances in the Structural and Mechanistic Aspects of Hsp70 Molecular Chaperones. *J. Biol. Chem.* **2019**, *294*, 2085–2097.
- (26) Scholl, Z. N.; Josephs, E. A.; Marszalek, P. E. Modular, Nondegenerate Polyprotein Scaffolds for Atomic Force Spectroscopy. *Biomacromolecules* **2016**, *17*, 2502–2505.
- (27) Scholl, Z. N.; Li, Q.; Marszalek, P. E. Single Molecule Mechanical Manipulation for Studying Biological Properties of

- Proteins, DNA, and Sugars. *WIREs Nanomed. Nanobiotechnol.* **2014**, *6*, 211–229.
- (28) Scholl, Z. N.; Yang, W.; Marszalek, P. E. Direct Observation of Multimer Stabilization in the Mechanical Unfolding Pathway of a Protein Undergoing Oligomerization. *ACS Nano* **2015**, *9*, 1189–1197.
- (29) Li, Q.; Scholl, Z. N.; Marszalek, P. E. Capturing the Mechanical Unfolding Pathway of a Large Protein with Coiled-Coil Probes. *Angew. Chem., Int. Ed.* **2014**, *53*, 13429–13433.
- (30) Meinhold, S.; Bauer, D.; Huber, J.; Merkel, U.; Weiß, A.; Žoldák, G.; Rief, M. An Active, Ligand-Responsive Pulling Geometry Reports on Internal Signaling between Subdomains of the DnaK Nucleotide-Binding Domain in Single-Molecule Mechanical Experiments. *Biochemistry* **2019**, *58*, 4744–4750.
- (31) Scholl, Z. N.; Yang, W.; Marszalek, P. E. Competing Pathways and Multiple Folding Nuclei in a Large Multidomain Protein, Luciferase. *Biophys. J.* **2017**, *112*, 1829–1840.
- (32) Kim, M.; Abdi, K.; Lee, G.; Rabbi, M.; Lee, W.; Yang, M.; Schofield, C. J.; Bennett, V.; Marszalek, P. E. Fast and Forceful Refolding of Stretched α -Helical Solenoid Proteins. *Biophys. J.* **2010**, *98*, 3086–3092.
- (33) Ganim, Z.; Rief, M. Mechanically Switching Single-Molecule Fluorescence of GFP by Unfolding and Refolding. *Proc. Natl. Acad. Sci. U.S.A.* **2017**, *114*, 11052–11056.
- (34) Dietz, H.; Rief, M. Protein Structure by Mechanical Triangulation. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 1244–1247.
- (35) Yu, H.; Siewny, M. G. W.; Edwards, D. T.; Sanders, A. W.; Perkins, T. T. Hidden Dynamics in the Unfolding of Individual Bacteriorhodopsin Proteins. *Science* **2017**, *355*, 945–950.
- (36) Borgia, M. B.; Borgia, A.; Best, R. B.; Steward, A.; Nettels, D.; Wunderlich, B.; Schuler, B.; Clarke, J. Single-Molecule Fluorescence Reveals Sequence-Specific Misfolding in Multidomain Proteins. *Nature* **2011**, *474*, 662–665.
- (37) Oberhauser, A. F.; Marszalek, P. E.; Carrion-Vazquez, M.; Fernandez, J. M. Single Protein Misfolding Events Captured by Atomic Force Microscopy. *Nat. Struct. Biol.* **1999**, *6*, 1025–1028.
- (38) Scholl, Z. N.; Li, Q.; Josephs, E.; Apostolidou, D.; Marszalek, P. E. Force Spectroscopy of Single Protein Molecules Using an Atomic Force Microscope. *J. Vis. Exp.* **2019**, *144*, No. e55989.
- (39) Florin, E. L.; Rief, M.; Lehmann, H.; Ludwig, M.; Dornmair, C.; Moy, V. T.; Gaub, H. E. Sensing Specific Molecular Interactions with the Atomic Force Microscope. *Biosens. Bioelectron.* **1995**, *10*, 895–901.
- (40) Bustamante, C.; Marko, J. F.; Siggia, E. D.; Smith, S. Entropic Elasticity of λ -Phage DNA. *Science* **1994**, *265*, 1599–1600.
- (41) Scott, D. W. Sturges' Rule. *Wiley Interdiscip. Rev. Comput. Stat.* **2009**, *1*, 303–306.
- (42) Noel, J. K.; Levi, M.; Raghunathan, M.; Lammert, H.; Hayes, R. L.; Onuchic, J. N.; Whitford, P. C. SMOG 2: A Versatile Software Package for Generating Structure-Based Models. *PLoS Comput. Biol.* **2016**, *12*, No. e1004794.
- (43) Clementi, C.; Nymeyer, H.; Onuchic, J. N. Topological and Energetic Factors: What Determines the Structural Details of the Transition State Ensemble and “en-Route” Intermediates for Protein Folding? An Investigation for Small Globular Proteins. *J. Mol. Biol.* **2000**, *298*, 937–953.
- (44) Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E. M.; Mittal, J.; Feig, M.; MacKerell, A. D. Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone ϕ , ψ and Side-Chain X1 and X2 Dihedral Angles. *J. Chem. Theory Comput.* **2012**, *8*, 3257–3273.
- (45) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926–935.
- (46) Bussi, G.; Donadio, D.; Parrinello, M. Canonical Sampling through Velocity Rescaling. *J. Chem. Phys.* **2007**, *126*, No. 014101.
- (47) Parrinello, M.; Rahman, A. Polymorphic Transitions in Single Crystals: A New Molecular Dynamics Method. *J. Appl. Phys.* **1981**, *52*, 7182–7190.
- (48) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An $N \log(N)$ Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98*, 10089–10092.
- (49) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. Gromacs: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers. *SoftwareX* **2015**, *1–2*, 19–25.
- (50) Lu, H.; Schulten, K. The Key Event in Force-Induced Unfolding of Titin's Immunoglobulin Domains. *Biophys. J.* **2000**, *79*, 51–65.
- (51) Perilla, J. R.; Schulten, K. Physical Properties of the HIV-1 Capsid from All-Atom Molecular Dynamics Simulations. *Nat. Commun.* **2017**, *8*, No. 15959.
- (52) Verdorfer, T.; Bernardi, R. C.; Meinhold, A.; Ott, W.; Luthey-Schulten, Z.; Nash, M. A.; Gaub, H. E. Combining in Vitro and in Silico Single-Molecule Force Spectroscopy to Characterize and Tune Cellulosomal Scaffoldin Mechanics. *J. Am. Chem. Soc.* **2017**, *139*, 17841–17852.
- (53) Sharma, D.; Perisic, O.; Peng, Q.; Cao, Y.; Lam, C.; Lu, H.; Li, H. Single-Molecule Force Spectroscopy Reveals a Mechanically Stable Protein Fold and the Rational Tuning of Its Mechanical Stability. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 9278–9283.
- (54) Milles, L. F.; Unterauer, E. M.; Nicolaus, T.; Gaub, H. E. Calcium Stabilizes the Strongest Protein Fold. *Nat. Commun.* **2018**, *9*, No. 4764.
- (55) Jagannathan, B.; Elms, P. J.; Bustamante, C.; Marqusee, S. Direct Observation of a Force-Induced Switch in the Anisotropic Mechanical Unfolding Pathway of a Protein. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 17820–17825.
- (56) Guinn, E. J.; Tian, P.; Shin, M.; Best, R. B.; Marqusee, S. A Small Single-Domain Protein Folds through the Same Pathway on and off the Ribosome. *Proc. Natl. Acad. Sci. U.S.A.* **2018**, *115*, 12206–12211.
- (57) Schwarz-Linek, U.; Werner, J. M.; Pickford, A. R.; Gurusiddappa, S.; Ewa, J. H. K.; Pilka, S.; Briggs, J. A. G.; Gough, T. S.; Höök, M.; Campbell, I. D.; Potts, J. R. Pathogenic Bacteria Attach to Human Fibronectin through a Tandem β -Zipper. *Nature* **2003**, *423*, 177–181.
- (58) Bornschlög, T.; Rief, M. Single Molecule Unzipping of Coiled Coils: Sequence Resolved Stability Profiles. *Phys. Rev. Lett.* **2006**, *96*, 118102.
- (59) Hyeon, C.; Dima, R. I.; Thirumalai, D. Pathways and Kinetic Barriers in Mechanical Unfolding and Refolding of RNA and Proteins. *Structure* **2006**, *14*, 1633–1645.
- (60) Carrion-Vazquez, M.; Oberhauser, A. F.; Fowler, S. B.; Marszalek, P. E.; Broedel, S. E.; Clarke, J.; Fernandez, J. M. Mechanical and Chemical Unfolding of a Single Protein: A Comparison. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 3694–3699.
- (61) Walder, R.; Van Patten, W. J.; Adhikari, A.; Perkins, T. T. Going Vertical to Improve the Accuracy of Atomic Force Microscopy Based Single-Molecule Force Spectroscopy. *ACS Nano* **2018**, *12*, 198–207.
- (62) Rivera, M.; Lee, W.; Ke, C.; Marszalek, P. E.; Cole, D. G.; Clark, R. L. Minimizing Pulling Geometry Errors in Atomic Force Microscope Single Molecule Force Spectroscopy. *Biophys. J.* **2008**, *95*, 3991–3998.
- (63) Tomabechi, Y.; Hosoya, T.; Ehara, H.; Sekine, S.; Shirouzu, M.; Inouye, S. Crystal Structure of NanoKAZ: The Mutated 19 KDa Component of Oplophorus Luciferase Catalyzing the Bioluminescent Reaction with Coelenterazine. *Biochem. Biophys. Res. Commun.* **2016**, *470*, 88–93.
- (64) Curto, L.; Angelani, C.; Delfino, J. Intervening in the β -Barrel Structure of Lipid Binding Proteins: Consequences on Folding, Ligand-Binding and Aggregation Propensity. *Prostaglandins, Leukotrienes Essent. Fatty Acids* **2015**, *93*, 37–43.
- (65) Zhang, P.; Wang, D.; Yang, W.; Marszalek, P. E. Piecewise All-Atom SMD Simulations Reveal Key Secondary Structures in Luciferase Unfolding Pathway. *Biophys. J.* **2020**, *119*, 2251–2261.

(66) Hohng, S.; Zhou, R.; Nahas, M. K.; Yu, J.; Schulten, K.; Lilley, D. M. J.; Ha, T. Fluorescence-Force Spectroscopy Maps Two-Dimensional Reaction Landscape of the Holliday Junction. *Science* **2007**, *318*, 279–283.

Recommended by ACS

Conformational Fluctuations in β 2-Microglobulin Using Markov State Modeling and Molecular Dynamics

Mahdi Ghorbani, Jeffery B. Klauda, *et al.*

AUGUST 01, 2023
THE JOURNAL OF PHYSICAL CHEMISTRY B

READ 

Structure Reveals the Impact of Surface Charge Distribution on the Phase Separation and Aggregation of Slr0280

Xiao-Dan Li, Kai-Hong Zhao, *et al.*

JUNE 15, 2023
BIOCHEMISTRY

READ 

Theory for High-Throughput Genetic Interaction Screening

Madeline E. McCarthy, Marc R. Birtwistle, *et al.*

JULY 18, 2023
ACS SYNTHETIC BIOLOGY

READ 

Heat-Induced Conformational Transition Mechanism of Heat Shock Factor 1 Investigated by Tryptophan Probe

Soichiro Kawagoe, Tomohide Saio, *et al.*

DECEMBER 09, 2022
BIOCHEMISTRY

READ 

Get More Suggestions >