Type III secretion system effector evolved to be mechanically labile and initiate unfolding from the N-terminus.

Katherine E. DaPron[ab], Alexandre M. Plastow[cd], Morgan R. Fink[a], Brandon Dzuba[e], Marc-Andre LeBlanc[ae] Thomas T. Perkins[be,1], Emad Tajkhorshid[cdf,1] and Marcelo C. Sousa[a,1]

[a]Department of Biochemistry, University of Colorado Boulder, Boulder, Colorado 80309
[b]Department of Molecular, Cellular and Developmental Biology, University of Colorado, Boulder, Colorado 80309
[c]Theoretical and Computational Biophysics Group, Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, IL 61801
[d]Center for Biophysics and Quantitative Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801
[e]JILA, National Institute of Standards and Technology and University of Colorado Boulder, Boulder, CO 80309
[f]NIH Center for Macromolecular Modeling and Bioinformatics, Department of Biochemistry, University of Illinois at Urbana-Champaign, Urbana, IL 61801

[1]To whom correspondence should be addressed:
Marcelo C. Sousa
Department of Biochemistry, 596 UCB
3415 Colorado Avenue
Boulder, CO 80309
E-mail: Marcelo.Sousa@colorado.edu

Thomas T. Perkins
JILA, 440 UCB,
University of Colorado, Boulder, Boulder, CO 80309
E-Mail: tperkins@colorado.edu

Emad Tajkhorshid
3147 Beckman Institute, MC-251
405 N. Mathews Ave.
Urbana, IL 61801
E-Mail: emad@illinois.edu

ORCID
Katherine E. DaPron: 0000-0002-0535-3898
Alexandre M. Plastow: 0009-0003-9726-2604
Morgan R. Fink: 0000-0002-0567-3234
Brandon Dzuba: 0009-0004-6478-0242
Marc-Andre LeBlanc: 0000-0002-5520-7389
Thomas T. Perkins: 0000-0003-4826-9490
Emad Tajkhorshid: 0000-0001-8434-1010
Marcelo C. Sousa: 0000-0002-0242-3619

Abstract

Many gram-negative pathogens critically depend on the Type III secretion system (T3SS) to inject effector proteins into host cells for colonization. Because the channel of the T3SS is narrow (~2 nm), effectors must be unfolded for secretion. However, the T3SS is unable to unfold mechanically robust substrates (GFP, ubiquitin, and dihydrofolate reductase), severely impairing their secretion. Consistent with this, effectors are exceptionally mechanically labile, unfolding at very low forces. Thus, secretion competency is correlated with mechanical properties. Effector sequences have significantly diverged from non-effectors despite shared folds, suggesting that secretion exerts evolutionary pressure selecting mechanical lability. Here, using atomic-force-microscopy–based force spectroscopy, we show that effector NleC is mechanically labile (F_{unfold} = 13.5 pN at 100 nm/s) and mechanically compliant, as characterized by a large distance to the transition state ($\Delta x^{\ddagger}$ = 2.7 nm). In contrast, the non-effector homolog protealysin is mechanically stable (F_{unfold} = 50.7 pN at 100 nm/s) and more brittle ($\Delta x^{\ddagger}$ = 0.7 nm), comparable to proteins known to impair secretion (F_{unfold} > 80 pN; $\Delta x^{\ddagger}$ < 0.4 nm). Denaturant-induced unfolding assays demonstrate that effectors exhibit rates typical of their fold, further reinforcing mechanical properties rather than intrinsically fast unfolding kinetics (k_0) predicts secretion. Steered molecular dynamic simulations revealed NleC unfolding initiates at the N-terminus, consistent with current secretion models, whereas protealysin unfolding initiates at the C-terminus. Notably, the NleC N-terminus is primarily α-helical while non-effector homologs contain β-sheets, which may account for the distinct unfolding pathway. Together, these results are consistent with mechanical lability being an evolved, structurally encoded feature underlying effector secretion.

Significance

The Type III secretion system delivers effector proteins directly into host cells to promote bacterial colonization. Effectors must be unfolded for secretion. This particular selective pressure is hypothesized to have driven significant sequence divergence from non-effector proteins despite maintaining the same fold. Here, we show that effectors are not characterized by unusually fast unfolding rates. Rather, effector NleC is more mechanically labile than its non-effector homolog, indicating mechanical lability underlies both effector sequence divergence and their unfolding. Simulations revealed that NleC unfolding initiates via the N-terminus consistent with the current secretion mechanism. Together, these results strongly suggest mechanical lability is an evolved property of effectors and provide structural insight into how it is encoded.

Introduction

The Type III secretion system (T3SS) is a specialized nanomachine of Gram-negative bacteria that delivers effector proteins directly into the host to facilitate colonization in symbiotic or pathogenic contexts. The T3SS complex resembles a nanoscale syringe that spans both the inner and outer membrane of the bacteria and host cell membranes, effectively connecting bacteria and host cytosol to inject effector proteins directly into the host (Fig. 1 *A*). Effectors are synthesized in the bacterial cytosol. They are characterized by an unstructured N-terminal T3SS targeting sequence, followed by a loosely structured chaperone binding domain for chaperone-dependent effectors and one or more well folded C-terminal effector domains (1-6). Effectors, alone or in their chaperone complexes, are trafficked to the base T3SS where their N-termini engage with the secretion apparatus and the T3SS ATPase catalyzes chaperone release (7). Because the inner diameter of the T3SS needle is only ~2 nm, secretion requires unfolding of effector domains. Current models propose that the effector N-terminus is threaded through the T3SS complex with subsequent unfolding and secretion powered by the ATPase and proton motive force (PMF) (8, 9). Effectors refold once translocated into the host cell and thereby modulate host cell processes to facilitate colonization such as downplaying the immune response (10-14) or inducing dramatic cytoskeleton rearrangements (15-17).

Effector unfolding is thus critical for secretion. In most cellular contexts, protein unfolding is catalyzed by AAA+ ATPases. These hexameric unfoldases can even unfold mechanically stable proteins (*e.g.*, GFP and ubiquitin)(18-21). However, the T3SS ATPase does not belong to the AAA+ family. Rather, it is homologous to the β subunit of the F_1 ATP synthase that operates as a rotary motor (22, 23). Therefore, the mechanism of T3SS effector unfolding remains poorly understood. It is nevertheless clear that, in contrast to AAA+ ATPases, the T3SS unfoldase cannot unfold certain substrates. Specifically, studies of effector fusions to GFP, ubiquitin, or dihydrofolate reductase (DHFR) show that secretion is strongly impaired as the fused domain resists unfolding (7, 24, 25). Structural analyses of these stalled complexes reveal that effectors are unfolded within the T3SS needle, whereas the folded fusion remains trapped in the basal body of the T3SS (26, 27). These findings establish that effector unfolding is a kinetic barrier that can be rate limiting due to the limited unfolding capacity of the T3SS unfoldase compared to canonical unfoldases (*e.g.*, ClpX). It has thus been proposed that T3SS effectors are thermodynamically unstable to accommodate their secretion (7, 24, 25, 28). However, recent work characterizing Salmonella effectors SopE2 and SptP showed that they have unremarkable thermodynamic stabilities, typical of globular proteins, but they are instead mechanically labile (29). This result has led to the new model that effectors are mechanically unfolded, and mechanical stability is a determinant for secretion as effectors are mechanically labile while substrates that are known to impair secretion (GFP, ubiquitin, and DHFR) are mechanically robust (29).

The requirement for effector proteins to be mechanically labile may be the origin of their extreme sequence divergence from non-effector homologs. Indeed, it typically requires three-dimensional structure analyses to identify non-effector homologs of T3SS effectors given their <15% sequence identity. Here, we directly test the hypothesis that effector proteins unfold at lower force than their non-effector homologs by comparing the mechanical stabilities of the T3SS effector NleC and its non-effector homolog protealysin. Both proteins are members of the MA clan of zinc-dependent proteases sharing a conserved

zincin fold. Using atomic force microscopy (AFM)–based single-molecule force spectroscopy, we show that NleC unfolds at low force and is therefore mechanically labile while protealysin is mechanically robust. Accompanying steered molecular dynamics simulations provided molecular insight by detailing unfolding pathways and demonstrated that NleC unfolds from the N-terminus when pulling across the N- and C-termini. Conversely, protealysin unfolds from the C-terminus indicating that its N-terminus is mechanically robust. Analysis of the NleC structure shows a divergent N-terminus in comparison with protealysin and other zincins consistent with the evolution of a mechanical labile structure to facilitate its unfolding by the T3SS unfoldase.

Results and Discussion

An effector and its Non-Effector Homolog: NleC and Protealysin.

To test our hypothesis that T3SS effector proteins may have evolved to be mechanically labile under selective pressure imposed by secretion, it was necessary to identify model effector and non-effector homolog proteins. Effectors that were mechanically characterized previously, SopE2 or SptP, do not have obvious non-effector homologs. The enterohemorrhagic *Escherichia coli* T3SS effector NleC, when delivered to the host cell, proteolytically cleaves the transcription factor NF-κB, thereby depressing transcription events mediating host immune responses (12-14, 30). NleC is a 332 amino acid protein with unstructured N-terminal 21 amino acids, rich in serine and proline residues, typical of the T3SS targeting signal of chaperone-independent effectors (1, 5, 6, 31, 32). NleC shares less than 15% sequence identity with any non-effector protein, further emphasizing the extreme sequence divergence of effectors. However, the crystal structure of NleC revealed that it belongs to the MA clan of metallopeptidases [MEROPS database (33)], with a basic Zincin fold of a three-helix bundle and mixed beta sheet around a conserved zinc-binding HExxH motif (31, 34, 35) (Fig. 1 *B*). Thus, a member of the MA metallopeptidase clan would be an appropriate reference to compare mechanical properties with NleC.

Thermolysin is perhaps the best characterized member of the MA metallopeptidases. However, it is expressed as a pre-proenzyme of 548 amino acids, and its larger size complicates the comparison with NleC. Nevertheless, protealysin was identified as a 341 amino acid thermolysin-like protease from *Serratia proteamaculans*, which does not have a T3SS, and is, thus, a suitable non-effector homolog of NleC. Like NleC, the structure of protealysin has been experimentally determined, revealing a well-structured propeptide and zincin fold (amino acids 6–341) although a short segment (amino acids 38–51) was not modeled due to conformational flexibility (36) (Fig. 1 *B*).

Effector NleC is Less Mechanical Stable than its Non-Effector Homolog Protealysin. We sought to compare the mechanical properties of NleC to protealysin using a previously developed AFM-based single-molecule force spectroscopy (SMFS) assay (29). In this assay, a polyprotein construct is stretched between a glass coverslip and an AFM cantilever (Fig. 1 *C*). The polyprotein incorporates an N-terminal ybbR-tag that is covalently attached to a PEG-CoA functionalized surface via the enzymatic activity of sfp (37). Following the ybbr-tag is a ddFLN4 marker domain, which unfolds with a characteristic two-step pattern and is used to screen for single-molecule attachment (38). Cohesin conjugated to PEG-coated cantilevers allows for site-specific attachment to the polyprotein via its mechanically robust interaction with dockerin (>300 pN) (39). This site-specific attachment to both a PEG-functionalized surface and cantilever minimized non-specific surface adhesion while the mechanically robust but reversible tip attachment enabled acquiring thousands of single-molecule force-extension curves with a single cantilever. Given the low unfolding force of effectors, we enhanced the force precision and time resolution by using focused-ion-beam-modified cantilevers (40-42).

Our single-molecule assay was initiated by gently pressing (75–100 pN) the cohesin-coated cantilever into the protein-coated surface for 0.15 s and then retracted the cantilever at a constant velocity (v). As the cantilever was retracted, the polyprotein was stretched until a protein domain unfolds, resulting in a rapid drop in force (F). Force was detected by cantilever deflection. The resulting force-extension curves (FECs) had a characteristic sawtooth pattern, where each peak corresponds to individual unfolding events (Fig. 1 *D*). For NleC, approximately 4% of the FECs were consistent with stretching a single polyprotein at v = 100, 400, 800, 1600, and 3200 nm/s, showing NleC unfolding followed by the characteristic double-peak structure of ddFLN4. (gray in Fig. 1 *D*). While Fig. 1 *D* shows a single unfolding intermediate for NleC, its unfolding pathway was heterogenous, showing 0 (8%), 1 (42%), 2 (37%) and 3 or more (13%) intermediates at 100-3200 nm/s (*SI Appendix,* Fig. S1). To analyze these curves, individual FEC segments were fit to a worm-like chain (WLC) model (43) (Fig. 1 *D*, dashed lines) that models the entropic stretching of a unstructured polymer with a persistence length (p = 0.365 nm (44) and contour length (L_0). For the NleC polyprotein, the ΔL_0 between the first peak in the FEC and the ddFLN4 marker domain was 91.6 ± 0.5 nm (mean ± SEM; N = 283) in good agreement with the expected value of 90.4 nm for NleC unfolding based on its structure and a contour length per amino acid of 0.365 nm/aa (44). NleC unfolded at low force [13.5 ± 0.5 pN (mean ± SEM; N = 64 at 100 nm/s)]. This is consistent with the previously reported mechanical stabilities of effectors SopE2 and SptP [16.5 ± 1.8 pN (mean ± SEM; N = 31) and 12.6 ± 0.7 pN (N = 25) at 100 nm/s, respectively], indicating that T3SS effectors are mechanically labile (29).

In pilot experiments, protealysin showed a significantly higher unfolding forces than NleC that were close in magnitude to the unfolding forces of the ddFLN4 marker domain. To facilitate analysis, we modified the polyprotein by substituting two GB1 domains, a more mechanically robust marker protein (45) for ddFLN4. With this new AFM construct, the resulting FEC yielded easily assignable peaks with protealysin unfolding before GB1 (Fig 1 *E*) with ~7% of cantilever retractions yielding single-molecule attachments at v = 30, 100, 300, 1000, and 3000 nm/s. Worm-like chain fits to the unfolding event showed that the ΔL_0 between the first unfolding peak and the first GB1 domain unfolding was 126.3 ± 0.7 nm (mean ± SEM; N =

156) good agreement with the expected value of 120.9 nm for unfolding protealysin. As with NleC, the FECs for protealysin frequently displayed additional unfolding peaks (pink in Fig. 1 *E*), consistent with 0, 1, 2 or 3 or more intermediates in 18, 48, 25 and 9% of the FECs from 30–3000 nm/s respectively (*SI Appendix*, Fig. S2).

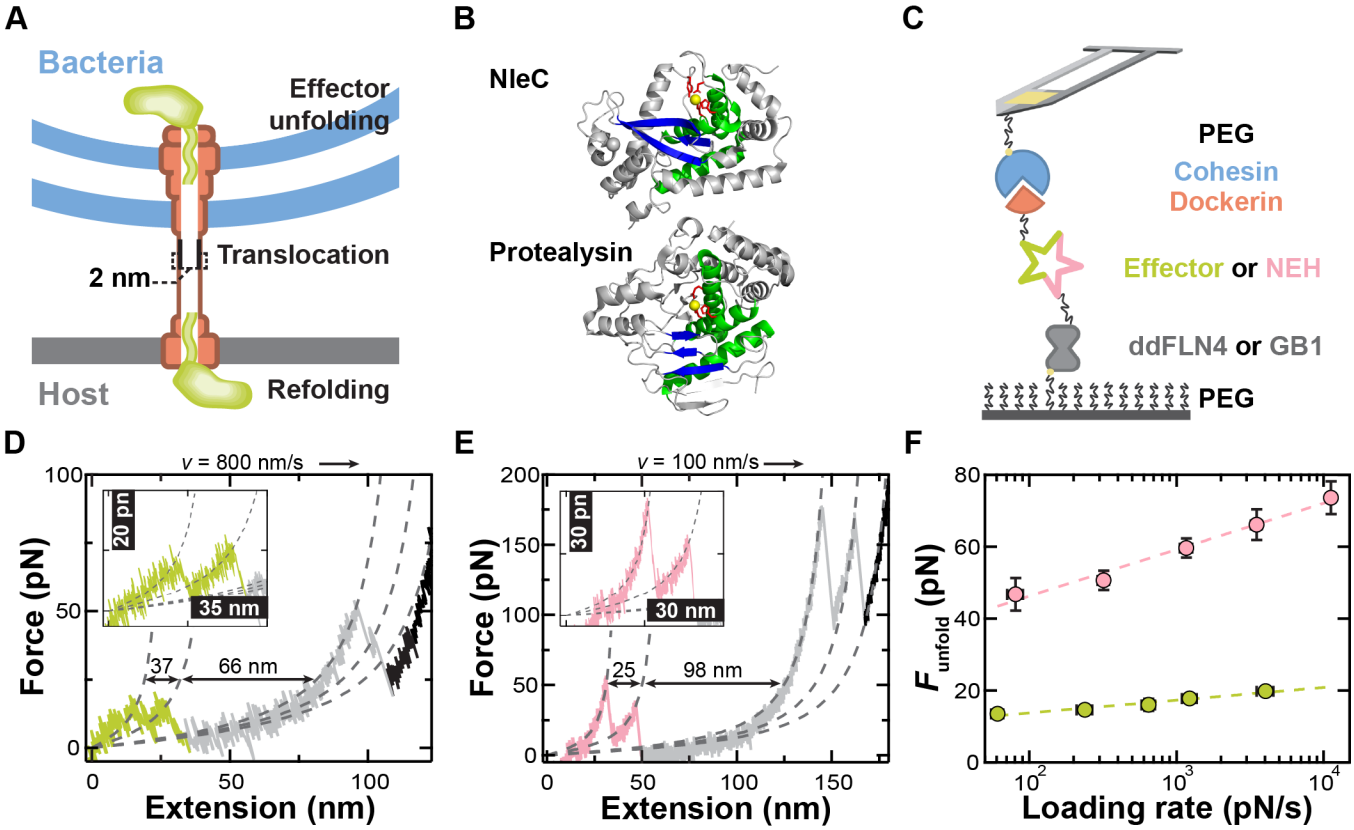

Figure 1 Mechanical stability of the Type III Secretion System (T3SS) effector NleC and non-effector homolog protealysin probed by a single-molecule atomic force microscopy (AFM) assay. (A) A schematic of Type III secretion. Effectors are first synthesized and folded in the bacterial cytosol. A T3SS unfoldase catalyzes effector unfolding, so effectors can translocate through the narrow needle of the complex (~2 nm). Effectors then need to refold in the host to modulate host cell processes. (B) Crystal structures of, NleC (top; PDB 4Q3J) and protealysin (bottom; PDB 2VQX), both zincin homologs, display highly conserved catalytic core centered on a zinc-binding motif, HExxH (shown in red sticks). This motif is composed of three helix bundle (green) and mixed beta sheet (blue) while peripheral details have diverged (grey). The catalytic zinc is shown in yellow. (C) Cartoon showing the single-molecule assay where retraction of a focused-ion-beam modified AFM cantilever stretches the polyprotein that is site specifically anchored to a PEG-coated surface. The polyprotein contains a well-defined marker protein (ddFLN4 or GB1, grey) followed by either the effector (green) or non-effector homologs (pink), and then dockerin (orange). Dockerin forms a strong but reversible protein-protein interaction with cohesin (blue) which enables stretching by a cohesin-coated AFM cantilever. (D,E) Representative force extension curves (FEC) at constant pulling velocity depict NleC (D, green) and protealysin (E, pink) unfolding as denoted by a sharp reduction in force. Segments of the FEC were fit with a worm-like chain model (dashed lines with the resulting change in contour length (ΔL_0) indicated by black arrows. Data smoothed to 5 kHz. (D) FEC was collected at 800 nm/s pulling velocity and depicts NleC unfolding (green) through a single intermediate followed by the characteristic two-step unfolding of ddFLN4 (grey). (E) FEC was collected at 100 nm/s pulling velocity and shows protealysin unfolding (pink) through a single intermediate and the unfolding of two copies of GB1 (grey). (F) The dynamic force spectra of NleC (green) and protealysin (pink) shows the average unfolding force versus the logarithm of the loading rate ($\partial F/\partial t$). Data point averages were determined from at least 20 unfolding events, and their error bars represent the standard error of the mean. Fitting the Bell-Evans model (dashed line) yielded the compliance, $\Delta x\ddagger$, and zero-force unfolding rate, k0.

To compare the mechanical properties of NleC and Protealysin, we measured their initial unfolding force over a range of velocities (30–3200 nm/s), yielding a dynamic force spectrum (46) (Figure 1F). In agreement with our hypothesis that effectors are mechanically labile compared to non-effector homologs, the average NleC unfolding force was more than 3-fold lower than protealysin over all velocities studied. More quantitatively, NleC unfolded at 13.5 ± 0.5 pN (mean ± SEM; N = 64 at 100 nm/s) while protealysin unfolded at 50.7 ± 2.7 pN (mean ± SEM; N = 37 at 100 nm/s), a highly significant difference ($p < 0.001$; unpaired T-test).

Effector NleC is Mechanically Compliant. Not only were the magnitudes of the unfolding forces between NleC and protealysin distinctly different, the slopes of the two dynamic force spectrums were as well (Figure 1F). This difference in slope for the average unfolding force plotted as a function of the logarithm of loading rate ($r = \partial F/\partial t$) arises from a change in the distance to the transition state ($\Delta x^{\ddagger}$), as described by Bell-Evans model (47). Within SMFS, $\Delta x^{\ddagger}$ can be interpreted as how much a protein deforms along the mechanical stretching reaction coordinate before unfolding. Thus, a large $\Delta x^{\ddagger}$ is associated with a compliant protein, whereas a protein with small $\Delta x^{\ddagger}$ is described as brittle. Of particular note for this study, GFP, ubiquitin and DHFR—all proteins that cannot be secreted by the T3SS—are brittle, with $\Delta x^{\ddagger}$ of 0.28, 0.23, and 0.37 nm, respectively (48-50).

Using the Bell-Evans model, we analyzed the dynamic force spectrum for NleC and protealysin, which yielded $\Delta x^{\ddagger}$ along with k_0, the zero-force unfolding rate. For NleC, $\Delta x^{\ddagger}$ was 2.7 ± 0.3 nm (mean ± fitting error), representing one the most compliant proteins measured to date by AFM (51). This result is consistent with the previously characterized effectors SopE2 and SptP that also proved to be highly mechanically compliant [$\Delta x^{\ddagger}$ = 1.6 ± 0.4 nm (mean ± fit error) and 1.4 ± 0.2 nm, respectively] (29). In contrast, analysis of the protealysin data yielded a $\Delta x^{\ddagger}$ = 0.7 ± 0.1 nm (mean ± fitting error)], showing the protealysin's compliance is significantly closer to proteins that inhibit secretion than its effector homolog.

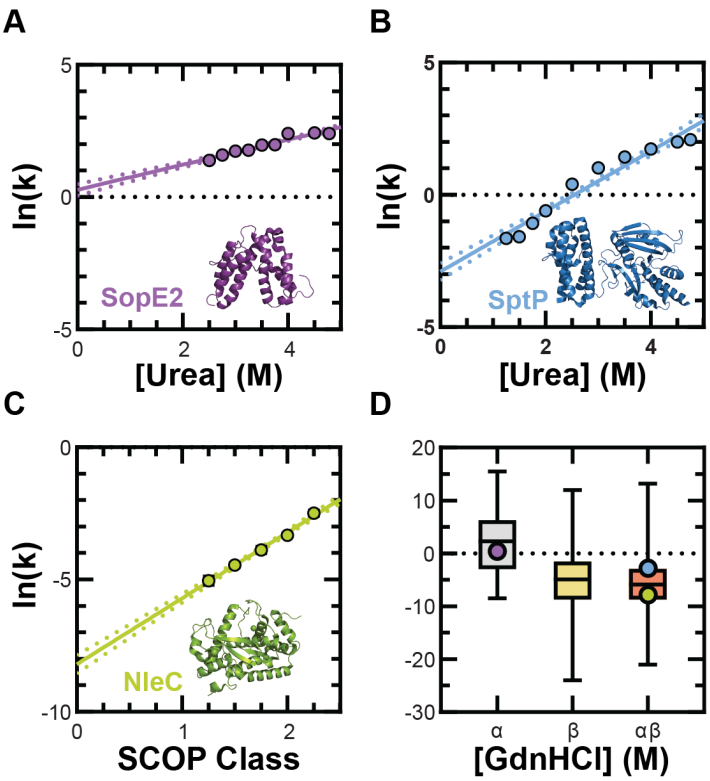

Figure 2 Kinetic unfolding of T3SS effector proteins. (A-C) The natural logarithm of SopE2 (A; purple), SptP (B, blue) and NleC (C, green) denaturant-dependent unfolding rates vary linearly with denaturant concentration, in which extrapolation to the absence of denaturant yields the natural logarithm of k0. For A-C, the average of three replicates was plotted, and the error bars represent the SD, but are smaller than the marker size. The data was fitted with Equation 3. Dotted lines represent the 95% confidence interval (D) Comparison of the natural logarithm of k_0 of T3SS effectors to reference proteins of various folds. K_0 boxplots are generated from N = 31, 37 and 40 (left to right) proteins categorized by SCOP class from reference (55). Boxes extend from 25th to 75th percentiles, and whiskers extend to minimum and maximum values. SopE2 (purple, all alpha-helical), SptP (blue, mixed topology), and NleC (green, mixed topology) are shown as individual values, representing the natural logarithm of their average k0. Error bars represent standard deviation; however, they are smaller than the marker size and therefore are not visible.

Compliance is relevant to effector secretion because there is a hyperbolic relationship between $\Delta x^{\ddagger}$ and unfolding forces for a set of mechanically characterized protein of diverse folds (51) (*SI Appendix,* Fig. S3). Previously, we showed that this hyperbolic relationship is described by the Bell-Evans model:

$$F_{mp} = (k_B T/\Delta x^{\ddagger})\ln[r\Delta x^{\ddagger}/(k_0 k_B T)] \qquad \text{Eq. 1}$$

where F_{MP} is the most probable unfolding force, r the loading rate and $k_B T$ is the thermal energy. This equation provides the scientific rationale for why a large $\Delta x^{\ddagger}$ leads to low force unfolding (29) and underscores the central role of $\Delta x^{\ddagger}$ in governing protein mechanical stability. The comparison with other proteins also highlights the selective pressure of secretion as NleC has usually low unfolding force and high compliance for a mixed topology fold whereas the mechanical properties of protealysin are more representative of its fold (*SI Appendix,* Fig. S3).

The large mechanical compliance of effectors bears physiological relevance. The Bell-Evans model defines unfolding rate under force (F) as:

$$k(F) = k_0 e^{F\Delta x^{\ddagger}/k_B T} \qquad \text{Eq. 2}$$

where k_0 is the intrinsic unfolding rate under zero-force. Therefore, for any given force, the unfolding rate depends exponentially on the $\Delta x^{\ddagger}$. The T3SS unfoldase is relatively weak and unable to unfold and secrete GFP, ubiquitin or DHFR. Nevertheless, even if small forces can be applied by the unfoldase, compliant proteins with large $\Delta x^{\ddagger}$, such as those in the effectors, will unfold faster than brittle proteins with short $\Delta x^{\ddagger}$. Accordingly, NleC is highly compliant while its non-effector homolog protealysin is quite brittle approaching the brittleness of GFP, ubiquitin and DHFR. This is consistent with NleC having evolved under pressure to be mechanically labile to facilitate efficient secretion by the TTSS.

Ensemble Unfolding Kinetics of SopE2, SptP and NleC Effector Domains. While the unfolding rate under force is exponentially dependent on the distance to the transition state ($\Delta x^{\ddagger}$), it is also linearly dependent on the intrinsic, zero-force unfolding rate (k_0) (Eq. 2). Therefore, it is possible that effectors may have evolved fast intrinsic unfolding kinetics, rather than mechanical lability, to facilitate their secretion. It was previously reported that zero-force unfolding rates (k_0) for SopE2 and SptP, as estimated from AFM data by fitting the Bell-Evans model to dynamic force spectra, did not correlate with secretion competency. However, these

estimates require extrapolation over orders or magnitude in loading rate, which is a known problem (52, 53). Therefore, we implemented an alternative approach of bulk-solution kinetics, measuring the unfolding kinetics of the SopE2, SptP and NleC effector domains by circular dichroism at increasing concentrations of denaturant. The natural logarithm of the observed exponential unfolding rate constants where then plotted vs the denaturant concentration and linearly fitted to extrapolate the unfolding rate constants in the absence of denaturant (Fig. 2 *A-C*). This yielded k_0 of (13.2 ± 1.1) x 10^{-1} s^{-1} (mean ± standard deviation [SD]), (55.2 ± 1.1) x 10^{-3} s^{-1} and (28.7 ± 1.1) x 10^{-5} s^{-1} for SopE2, SptP and NleC, respectively.

We then compared these observed unfolding rate constants to those in a set of 108 proteins with experimentally determined unfolding kinetics spanning all alpha, all beta or mixed alpha beta structural classes (54, 55). As shown in Fig. 2 *D*, the unfolding rates for the all-alpha SopE2 effector or the mixed alpha-beta SptP and NleC effectors are not significantly different than those of reference proteins in the same structural class. Therefore, we conclude that effector proteins do not have unusually fast intrinsic unfolding kinetics that may explain their secretion competency.

Steered Molecular Dynamics Simulations Reveal the NleC and Protealysin Pathways of Unfolding Under Force. The molecular events associated with protein unfolding in SMFS experiments can be interpreted with help from all-atom steered molecular dynamics (SMD) simulations (39, 56-64). In this approach, the protein atomic model is anchored at one end while the other end is pulled using a harmonic restraint moving at a constant velocity (constant velocity SMD). These simulations produce trajectories that associate specific unfolding transitions with the unfolding forces (56) and provide insight into the secondary and tertiary structural features responsible for mechanical stability.

We carried out three replicas of all-atom SMD simulations of NleC (amino acids 22-280). To reproduce the pulling geometry of the AFM experiments, the N-terminus was harmonically restrained to its initial position, and the C-terminal residue of the protein was pulled at 2.064 Å/ns. The applied force on the harmonic spring was monitored and plotted versus the distance from the initial position to generate a simulated force-extension trace (Fig. 3 *A*).

All three replicas showed similar unfolding pathways (*SI Appendix*, Fig. S4 A-C) where the first unfolding event was always detachment of the N-terminal helix. This corresponded with the first peak and force drop in the simulated FEC (Fig. 3 *AB* and *SI Appendix,* Fig. S5 AB, events labeled 1 and 2). Unfolding continued from the C-terminal end of the protein (Fig. 3 *C*)

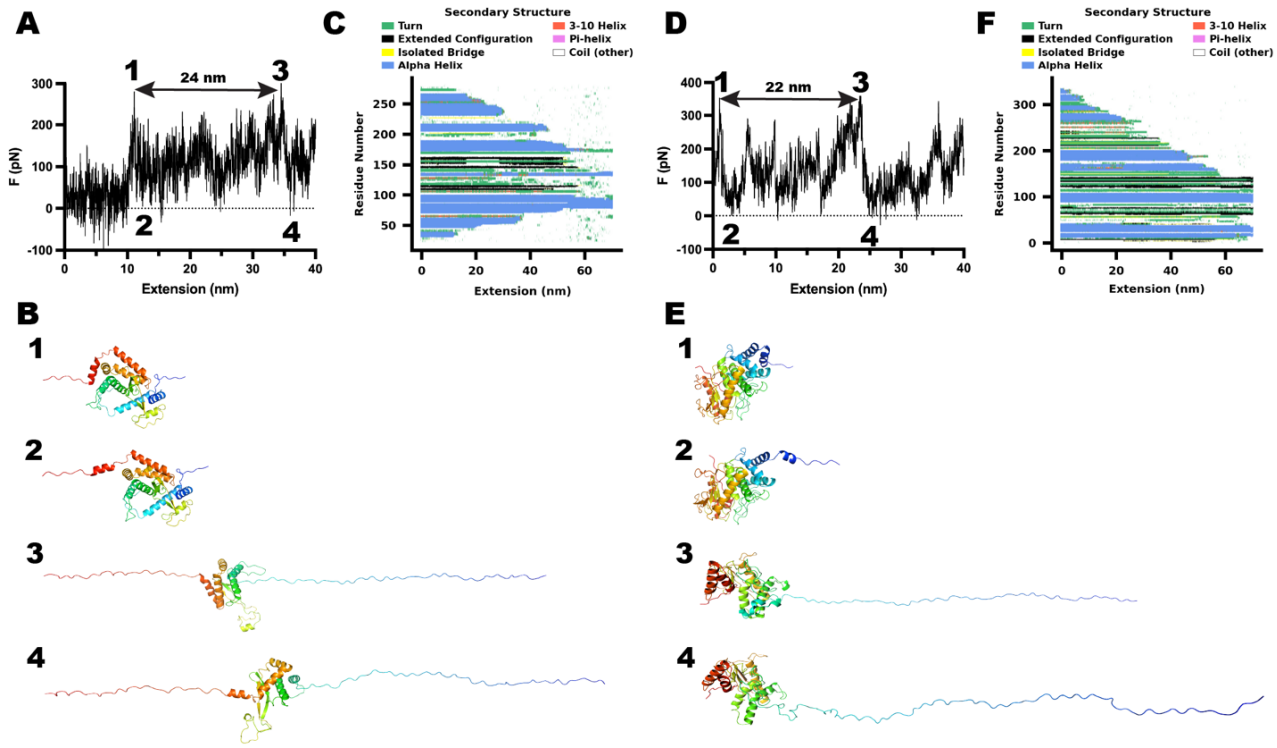

Figure 3 Force induced unfolding behavior of NleC and protealysin in SMD simulations. (A and D) Representative simulated force extension curve for NleC (A) and protealysin (D) showing the first 40 nm of extension. The major force peaks (1 and 3) and force drops (2 and 4) are indicated. FEC was smoothed with the Savitzky–Golay filter for visual presentation. (B and E) Secondary structure evolution plot of the SMD simulation unfolding of NleC (B) and protealysin (E). Secondary structure elements in each simulation frame were assigned using DSSP and colored as described in the color key (C and F) Cartoon representations of the atomic models in the SMD simulation frames of NleC (C) and protealysin (F) unfolding corresponding to events 1-4 indicated in (A) and (D) respectively. The proteins are colored in a spectrum from red (N-terminus) to orange to yellow to green to cyan to blue (C-terminus) to indicate directionality.

until ~34nm of extension. At this point, the second N-terminal helix is detached and unfolded corresponding to another peak and force drop in the FEC similar in magnitude to the first peak (Fig. 3 *AB* and *SI Appendix,* Fig. S5 AB, events labeled 3 and 4). These two peaks were separated by 24 nm in extension, which is close to the ΔL_0 of one of the most common intermediates observed in the AFM experiments (~27 nm) (*SI Appendix*, Fig. S6). We note that extension approximately equals L_0 at the elevated rupture forces in the SMD trajectories (>200 pN). Energy barriers in unfolding pathways can be broad (65) and crossing them is a stochastic, thermally driven process. Therefore, the height and precise position of peaks in the FEC varies from trajectory to trajectory in AFM experiments and SMD simulations. Nevertheless, there is broad agreement between the experimental ΔL_0 and the extension difference between peaks in the simulated FECs (*SI Appendix*, Fig. S6). This indicates that the SMD simulations are a useful tool to identify the molecular events associated with unfolding in the AFM experiments. For NleC, initial unfolding occurs at the N-terminus, and it required low forces in the AFM experiments (Fig. 1). All subsequent unfolding events, from either the N- or C-terminus, required forces similar to the initial peak in both the AFM and SMD force-extension curves (Fig. 1 and 3).

All atom SMD simulations were also carried out for protealysin (amino acids 1-341) at a pulling speed of 2.73 Å/ns in triplicate. All simulations showed a similar unfolding pathway (*SI Appendix*, Fig. S4 D-F). In stark contrast to NleC however, the first unfolding event was always detachment of the C-terminal helix which corresponds to a peak and force drop in the simulated FEC (Fig. 3 *DE* and *SI Appendix,* Fig. S5 CD, events labeled 1 and 2). Protealysin then unfolded sequentially from the C-terminus (Fig. 3 *F*). A broad force peak and force drop similar in magnitude to the initial peak was observed at approximately 24 nm extension, which corresponds to detachment and unfolding of a helix-turn-helix motif that is C-terminal of the conserved zincin fold of protealysin (Fig. 3 *DE* and *SI Appendix,* S5 CD, events labeled 3 and 4). The peaks are separated by 22 nm in extension, in good agreement with the change in contour length for the most common intermediate observed in the AFM experiments (ΔL_0 ~22 nm) (*SI Appendix*, Fig. S7). This agreement supports the usefulness of the SMD trajectories to interpret the AFM unfolding experiments. For protealysin, unfolding proceeds from the C-terminus which required higher forces than NleC in the AFM experiments. This implies that the N-terminal domain of protealysin is mechanically robust compared to its C-terminus and that detaching it from the intact structure would require forces even higher than those observed in the AFM experiment.

Biological Implications. T3SS effectors are synthesized and folded in the bacterial cytosol prior to secretion. Some effectors, such as SopE2 and SptP, have an N-terminal, approximately 100 amino acid long, chaperone binding domain. Specialized chaperones bind these regions and keep them in extended, loosely folded conformations, whereas the C-terminal effector domains are well folded even when in complex with their cognate chaperones (66-68). Similarly, chaperone-independent effectors such as NleC are well folded in the bacterial cytosol. Indeed, expression of *Citrobacter rodentium* NleC together with its eukaryotic substrate NF-κB p65 resulted in intra-bacterial cleavage of p65 demonstrating that NleC was enzymatically active and thus folded (69).

As the lumen of the T3SS needle is approximately 2 nm in diameter, effector proteins must be unfolded for secretion through the T3SS. However, the T3SS unfoldase is relatively weak and unable to unfold mechanically stable proteins such as GFP, ubiquitin and DHFR. Previous work established that effector proteins have typical thermodynamic stabilities, and thus thermodynamic stability does not correlate with secretion competency (29). Conversely, mechanical stability was negatively correlated with secretion competency. Specifically, *Salmonella* effectors SopE2 and SptP were shown to be mechanically labile, unfolding at low forces, and compliant, with long distance to the transition state, while GFP, ubiquitin and DHFR were mechanically robust, unfolding at higher forces, and brittle, with short distances to the transition state (29).

Here we showed that the enterohemorrhagic *E. coli* effector NleC is also mechanically labile and unusually compliant, further cementing these features as the mechanical fingerprint that defines T3SS effectors. NleC belongs to the MA clan of zinc-dependent proteases containing a conserved zincin fold. Notably, the NleC non-effector homolog protealysin, another MA clan zincin, was more mechanically stable requiring higher forces for unfolding, as well as brittle, with a short distance to the transition state. These observations support a model where the effector NleC diverged from other zincins under pressure to evolve structural features that would confer mechanical lability for efficient unfolding by the weak T3SS unfoldase. Such a particular evolutionary pressure is consistent with the sequence divergence observed in T3SS effectors, which in the case of NleC, results in less than 15% sequence identity with any other non-effector zincin.

SMD simulations of NleC and protealysin provide molecular insight into the unfolding pathways under force. The pulling geometry in these simulations across the N- and C-termini of the proteins is identical to that in the AFM experiments. Under these conditions, NleC consistently unfolded from the N-terminus, indicating that this is the part of the protein easiest to unfold by force. This is relevant for the T3SS unfoldase that may work by pulling effectors from their N-termini through a narrow pore to force their unfolding. For NleC, such an unfoldase would unfold the protein from its weakest point. Conversely, protealysin consistently unfolded sequentially from its C-terminus implying that its N-terminus is more resistant to unfolding. In the AFM geometry, protealysin required higher forces than NleC to be unfolded, but even higher forces would be required to unfold protealysin from its N-terminus T3SS unfoldase during secretion. This observation leads to the hypothesis that NleC evolved an N-terminal structure divergent from other zincins to be mechanically labile. Fig. 4 shows the three-dimensional structures of NleC and seven representative zincins with the conserved zincin fold colored green and blue and the segment N-terminal to the conserved fold colored red. Notably, NleC has two N-terminal helices that pack against the conserved zincin fold helices and three stranded β-sheet. On the other hand, all other zincins have N-terminal structures that add one or more β-strands to the conserved β-sheets in the zincin fold. As α-helices are mechanically labile

compared to β-sheets, this is consistent with the model that NleC evolved a divergent N-terminal structure that is mechanically labile to facilitate unfolding by a relatively weak T3SS unfoldase that pulls it substrates from their N-terminal end.

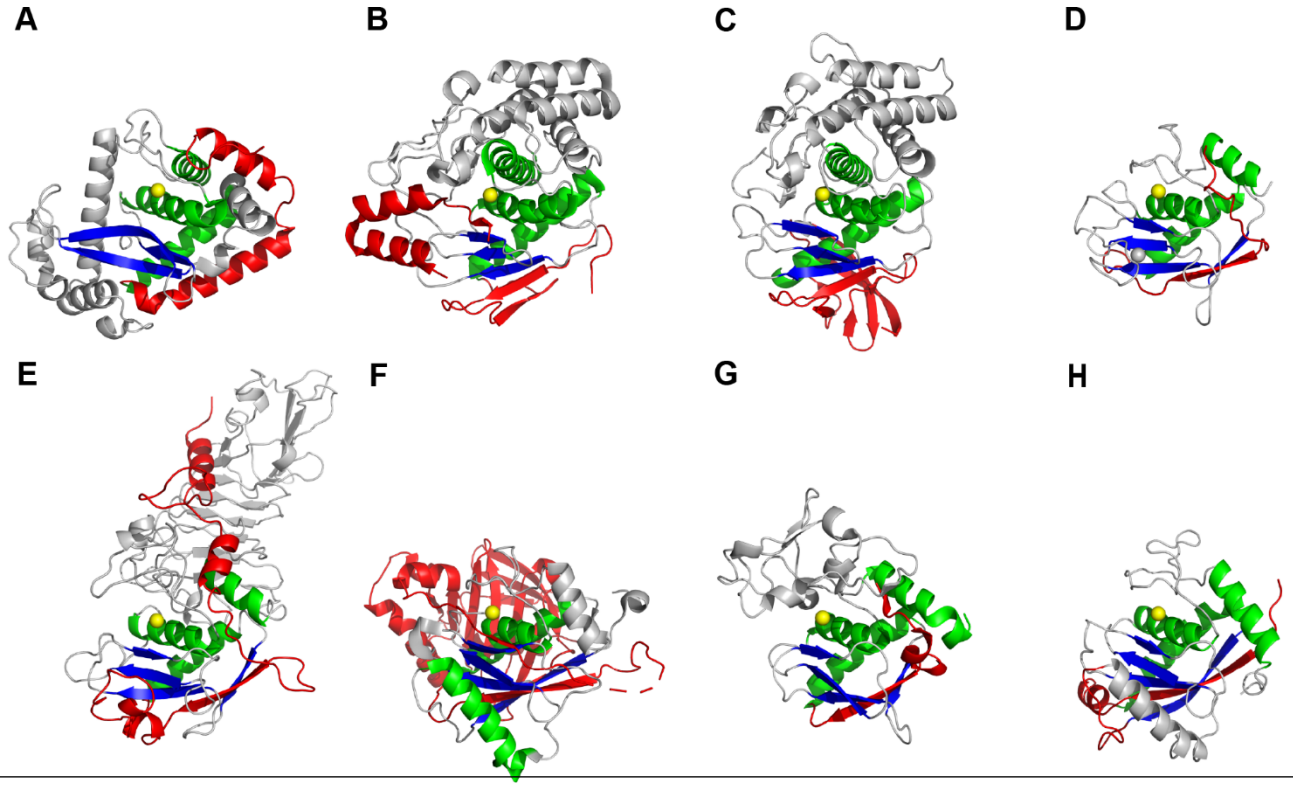

Figure 4 Divergent N-terminal region of representative Zincins. Zincins' (A-H) folds (grey) are defined by a conserved three helix bundle (green) and beta sheet (blue) surrounding the catalytic zinc (yellow) and zinc binding motif. The N-terminal peripheral structures (red) to the catalytic core are divergent. (A) NleC (M85 family; PDB 4Q3J); (B) protealysin (M4 family; PDB 2VQX); (C) thermolysin (M4 family; PDB 1TLP); (D) matrix metalloprotease-8 (M10A family; PDB 1KBC); (E) aeruginolysin (M10B family; PDB 1AKL); (F) fragilysin (M10C Family; PDB 3P24); (G) astacin (M12 family; PDB 1AST); (H) astrolysin (M12 family; PDB 1DTH) .

Materials and Methods

AFM constructs and cloning

The gene encoding the catalytic domain NleC (residues 22–280) was amplified by PCR using primers U2418 and L2418 (*SI Appendix*, Table S1). The gene encoding the full-length inactive protealysin (residues 6–341, E163A) was amplified by PCR using primers U2515 and L2483. NleC gene fragments were each incorporated into plasmid pMS-1480 linearized by PCR amplification with primers U2414 and L2414 by Gibson assembly using NEBuilder HiFi DNA Assembly (New England Biolabs). The protealysin gene fragments were integrated into plasmid pMS-1711 linearized by restriction digest with ZraI also by Gibson assembly. The construct with NleC is pMS-1620 and protealysin is pMS-1696. The vectors pMS-1480 and pMS-1711 are both pET-28a based expression vectors. pMS-1480 includes a 6x His tag for Ni-NTA affinity chromatography, a ybbR-tag that allows for covalent attachment to Coenzyme A (CoA) derivatized surfaces, well characterized 4th domain of the F-actin cross-linking filamin rod of Dictostelium discoideum (ddFLN4) as a marker protein (38) for single-molecule attachment, and Xmod-Dockerin for strong, noncovalent attachment of the polyprotein to Cohesin functionalized cantilevers. pMS-1711 has all the same features as pMS-1480 except it includes a different marker protein, streptococcal B1 immunoglobulin-binding domain of protein G or GB1 (45), rather than ddFLN4.

Protein expression and purification for AFM

The plasmids encoding the polyproteins were transformed into E. coli BL21 (D3E) and single colonies were grown overnight at 37 °C with shaking in LB supplemented with kanamycin. One liter LB+kanamycin cultures were inoculated (1:100 v/v) and grown at 37 °C with shaking. Once an OD600 of 0.6 was reached, cultures were put on ice for 30 min and then induced with 0.4 mM IPTG at 16 °C overnight.

Cultures were pelleted at 4000 rpm with a Beckman JLA 8.100 rotor for 30 min at 4 °C, and cell pellets were resuspended with lysis buffer composed of 50 mM Tris pH 8 supplemented with 0.5 mg/mL lysozyme and EDTA-free protease inhibitor tablets (Pierce). Cells were lysed by sonication on ice. Cell debris was removed by centrifugation in a Sorvall SS-34 rotor at 14,000xg for 30 min at 4 °C. The supernatant was filtered through 0.22 µM PVDF membrane before being loaded onto a

5 mL Ni-NTA (Cytiva) column equilibrated with Buffer A (50 mM Tris pH 8, 150 mM NaCl). The column was then washed with 5 CVs of Buffer A + 20 mM Imidazole and eluted off the column with a linear gradient of 5 CVs of Buffer A + 25-300 mM Imidazole. Fractions containing the protein of interest were then loaded onto a size exclusion column (HiLoad 26/60 Superdex 200, Amersham Pharmacia Biotech) equilibrated with AFM buffer (25 mM HEPES pH 7.2, 150 mM NaCl, 0.5 mM TCEP). Fractions containing the protein of interest were combined, concentrated, and stored at -80 °C for later use.

Functionalization of Cantilevers and Surfaces for AFM

Cantilevers (BioLever Mini, Olympus) were modified with a focused-ion beam (FIB) according to previously described methods (40, 70). Briefly, cantilevers were modified to remove large sections to reduce hydrodynamic drag, thinned to reduce stiffness, and chemically etched to remove the majority of the gold and chromium coatings to improve stability. We initially cleaned glass surfaces (German Glass) for functionalization by rinsing in subsequent baths of ultrapure water, acetone, ethanol, 3M KOH, and two additional rinses with ultrapure water. Surfaces were then dried with N2 gas and irradiated with ozone for 30 min (PSD-UV8, Novascan Technologies). Cantilevers and surfaces were then silanized and PEGylated using one of two methods described below.

In the first approach, maleimide-functionalized surfaces and cantilevers were created using established methods (71). Briefly, surfaces and cantilevers were reacted with a silane-PEG-maleimide (SPM, PEG MW=1000 Da, Nanocs, 0.15 mg/mL in toluene) at 65 °C for 3 h. After the incubation, we washed the cantilevers and surfaces sequentially in toluene, isopropanol, and two ultrapure water baths.

The second approach involved creating amine-functionalized surfaces and cantilevers prior to PEGylation. Surfaces were incubated with 3-aminopropyldimethylethoxysilane [APDMES; 1.8% (v/v)] in ethanol) for 1 h at room temperature whereas cantilevers were incubated with APDMES [50% (v/v)] for 15 min at room temperature. Surfaces and cantilevers were then washed sequentially in isopropanol and ultrapure water, dried, and incubated in an oven under vacuum at 80 °C for 1 h. Amine-functionalized cantilevers and surfaces then proceeded to PEGylation where they were incubated with borate buffer (50 mM Na2B4O7, pH 8.5) for 1 h at room temperature. A solution of maleimide-PEG12-TFP ester [25 mM in 25% DMSO (v/v): 75% (v/v) borate buffer] was made and centrifuged for 10 min at 15,000 x g. After removing excess buffer, the supernatant of the maleimide-PEG12-TFP ester solution reacted with surfaces or cantilevers reacted for 1 h at room temperature. They were then rinsed with ultrapure water three times.

We immediately followed PEGylation by chemically coupling the maleimide to CoA to minimize maleimide group hydrolysis. Surfaces and cantilevers were reacted with a CoA solution (1 mM in 50 mM sodium phosphate pH 7.2, 50 mM NaCl, 10 mM ethylenediaminetetraacetic acid) for 1 h at room temperature or overnight at 4 °C. We coupled the polyproteins to CoA surfaces by Sfp phosphopantetheinyl transferase (sfp), which covalently attaches CoA to the ybbR-tag encoded in our polyprotein constructs (37). AFM polyproteins were thawed and diluted to 0.1-1 μM in AFM buffer (25 mM HEPES, pH 7.2, and 150 mM NaCl) and combined with final concentrations of 3 μM SFP and 10 μM MgCl2 and incubated for 1 h at room temperature or overnight at 4 °C. Cohesin was also deposited onto CoA-functionalized cantilevers using the same procedure with final concentrations of 20 μM cohesion, 3 μM SFP and 10 μM MgCl2. The surfaces and cantilevers were then thoroughly rinsed with AFM buffer and then immediately used for AFM measurements.

AFM Assay and Analysis

We used a Cypher ES (Asylum Research) outfitted with a custom small spot laser (41) and temperature-controlled (25 ºC) fluid cell to collect AFM data sets. Cantilever sensitivity (V/nm) was determined by hard contact with a rigid surface and the spring constant (k) via the thermal method (72). Cantilever spring constants (k) ranged from ~6–15 pN/nm. To attach polyprotein to the cantilever, the cantilever was gently (75–100 pN) pressed into the surface for 0.15 s. The cantilever was then retracted at a constant velocity ranging from 30–3200 nm/s. A 150-nm spaced grid was probed across the surface to find and measure new molecules. The resulting traces were digitized at 50 kHz with a low-pass filter at 25 kHz. If the cantilever stopped yielding attachments, the cantilevers were cleaned and unfolded in urea bath (8M in AFM buffer) for 30 min. Cohesin on the cantilever was then refolded by soaking in AFM buffer for 30 min. Sensitivity and cantilever spring constant were recalibrated prior to use. This cleaning procedure often rescued attachment of the cantilever to the polyprotein. AFM experiments were stopped after ~3,000–10,000 force extension curves (FECs) were collected.

For the resulting FECs, force was determined from the deflection of the cantilever using the calibrated cantilever spring constant and sensitivity. Extension was determined by stage movement minus cantilever deflection. We calculated the loading rate and rupture force by fitting a line to a force-vs-time plot immediately before an unfolding event. In a small percentage of traces, the zero-force baseline was corrected to account for experimental drift. The FEC traces were then smoothed by box-car averaging to reduce noise for visual presentation. All peaks were fitted by the worm-like chain (WLC) model to determine the contour length (L_0) of peaks. Only force extension curves displaying change in contour lengths (ΔL_0) within 20% of expected values (GB1 = 18 nm; ddFLN4 = 32.2 nm; NleC = 90.4 nm; protealysin = 120.9 nm) were included in analyses. The unfolding force for NleC and protealysin was determined from the first unfolding event in the FEC when unfolding intermediates were observed. Rare high force unfolding forces exceeding 2 standard deviations from the mean were excluded from analyses, as they were likely due to cantilever-surface adhesion or protein misfolding.

CD constructs and cloning

The catalytic domain of NleC (residues 22–280) was amplified using primers U2442 and L2438 (*SI Appendix*, Table S1). pMS-984 was digested with XhoI and SfoI. The PCR product was digested with XfoI and ligated into digested pMS-984.

The plasmids encoding the catalytic domains of SopE2 (residues 69–240), SptP (residues 161–543) were obtained from reference (29). The constructs encoding SopE2, SptP and NleC are as follows: pMS-1599, pMS-1597, and pMS-1623, respectively. pMS-984, a pET-28-a based expression vector, encodes a N-terminal 6x His tag for Ni-NTA affinity purification, a N-terminal cleavable SUMO tag to increase protein solubility, and a multiple cloning site as an alternative method to clone genes of interest.

Protein expression and purification for CD Studies

The CD constructs were expressed and purified in a similar manner. CD constructs were transformed into E. coli BL-21 (DE3) cells for overexpression. Overnight cultures inoculated 1-liter cultures (1:100 v/v) of LB medium supplemented with Kanamycin. Cultures grew at 37 °C with shaking until OD600 0.6. Cells were then put on ice for 30 min and then induced with 0.4 mM IPTG. Cells were grown overnight at 16 °C and then pelleted for 30 min at 4 °C in a Beckman JLA 8.100 rotor at 4000 RPM.

Cell pellets were resuspended with lysis buffer [50 mM Tris pH 8, 0.5 mg/ml lysozyme, and EDTA-free protease inhibitor tablets (Pierce)]. The cells were lysed with a sonicator on ice. The cell lysate was separated from cell debris by centrifugation with a Sorvall SS-34 rotor at 14,000g for 30 min at 4 °C. The soluble protein fraction was loaded into a 5 mL Ni-NTA column (Cytiva) equilibrated with buffer A (50 mM Tris pH 8, 150 mM NaCl). The column was washed with Buffer A + 20 mM imidazole for 5 column volumes (CVs). The column was eluted with a linear gradient of Buffer A + 20-300 mM imidazole. The eluted fractions with protein were combined and 1:1000 (w/w) of His-tagged Ulp1 enzyme was added to cleave off the N-terminal SUMO and His tag. Eluted protein and His-tagged Ulp1 were dialyzed against the reaction buffer (25 mM Tris pH 8, 150 mM NaCl, 5 mM BME) with a 6-8 kDa membrane cutoff. The cleaved protein was then loaded onto a 5 mL Ni-NTA column using the same protocol except the flow-through fractions were collected as the eluted fractions contain His-tagged Ulp1, His-tagged SUMO-tag, and uncleaved protein. The cleaved protein was then loaded onto a Superdex75 column equilibrated with buffer A. The fractions containing only protein were combined and concentrated. NleC, SptP and SopE2 for CD were stored at -80 °C.

Kinetic Unfolding Measurements and Analysis

Protein stocks were thawed from -80 °C storage and filtered with a 0.1 µM PVDF filter. A 10M urea buffered (SopE2: 25 mM HEPES pH 7.2, 150 mM NaCl; SptP: 10 mM Tris pH 8, 150 mM sodium Sulfate) solution was made and deionized with BioRad AG 501-X8 resin (50 g/L) for 1 h. The resin was removed by vacuum filtration. A 6M guanidine hydrochloride buffered (25 mM HEPES pH 7.2, 150 mM sodium sulfate) solution was made. The concentrations of the urea and guanidine hydrochloride stocks were confirmed with an Abbe refractometer.

Stopped flow data for SptP and SopE2 was collected with an Applied PhotoPhysics ChiraScan Plus outfitted with a SX-20 Stop Flow spectrophotometer with a 20 µL flow cell. Before use, the spectrophotometer was thoroughly purged with nitrogen to prevent ozone formation and protect the optics. Multiple shots (5–10) of water and buffer completely purged the lines between measurements. CD absorbance spectra confirmed the system was equilibrated. Unfolding was initiated by the rapid mixing of 1:1 shot of protein and buffered urea at various concentrations chosen based on the concentration range of the transition region (29) with a dead time of 1.2 ms. The final concentration of protein after mixing was 0.18 mg/mL for SopE2 and 0.1 mg/mL for SptP. Kinetic data was collected for 5–120 s, collecting 1000 times points, at 222 nm. The collection time varied so that there would be good coverage of data points in the initial region of the unfolding curve. Three independent replicates were collected for each urea concentration. The replicates were averaged and fitted with a single or double exponential fit to extract k-fast rates of unfolding for each urea concentration.

The natural logarithm of the rate of unfolding linearly varies with [D]:

$$\ln[k]_D = \ln[k]_{H20} + m[D] \qquad \text{Equation 3}$$

where $[k]_D$ is the rate of unfolding in denaturant; $[k]_{H20}$ is the rate of unfolding in water or intrinsic unfolding rate, m is the constant of proportionality; and [D] is denaturant. The natural logarithm of the k-fasts were plotted against urea concentrations and fitted with equation 3 and 95% confidence intervals. The fit was extrapolated to the y-intercept to determine k_0.

Observing NleC unfolding kinetics required guanidine hydrochloride concentrations that could not be achieved in stopped flow due to 1:1 dilution. Thus, we opted for manual mixing which allowed higher guanidine hydrochloride concentrations that still captured initial unfolding rates despite increased experimental dead time. For manual mixing unfolding experiments with NleC, a quartz cuvette (Hellma) with a 10-mM path length was used with an Applied PhotoPhysics ChiraScan Plus spectrophotometer. Before use, the spectrophotometer was thoroughly purged with nitrogen to prevent ozone formation and protect the optics. NleC was mixed in the cuvette with the buffered guanidine hydrochloride to a final concentration of 0.05 mg/mL. The time from mixing to the beginning of the data collection was noted to correct for time (~15 s). Kinetic data was collected for 120–300 s, collecting 1000 times points logarithmically at 225 nm. The cuvette was cleaned with three 10M urea washes, three buffer washes, 1–2 1% Hellmanex solution washes, five ultrapure water washes, and then one 100% ethanol wash and house air to dry the cuvette in between measurements. The resulting denaturation curves were analyzed in a similar manner using previously described procedures.

SMD simulations

SMD simulations (73, 74) were performed to describe the unfolding mechanism of NleC and protealysin. The crystal structure of NleC (PDB: 4Q3J) and an AlphaFold3 model of protealysin (derived from PDB: 2VQX) were used as the starting structures for the simulations. Protein topologies were generated using VMD's psfgen plugin (75). The proteins were solvated and ionized using VMD's solvation and autoionize plugs (75). The net charge of the system was neutralized using a 150 mM of NaCl. To relax the structure and prepare for SMD simulations, the proteins were minimized for 10,000 steps and equilibrated for 5ns with a harmonic constrain with a force constant (k) of 1 kcal/mol/Å^2. SMD simulations were performed once the protein structure was properly equilibrated and checked for stability. The solvent box was enlarged in the Z direction to accommodate at least 300,000 atoms, which allows for a complete unfolding of a 180 amino-acid protein. Multiple SMD simulations were performed at pulling velocities of 2.06 Å/ns (NleC) and 2.73 Å/ns (protealysin) using the end-to-end distance of the protein as the reaction coordinate. In all the simulations the N-terminal residue of the protein was harmonically restrained to its initial position and a moving restraint with k = 10 kcal/mol/Å^2 was applied to the C-terminal residue of the protein. The applied force on the harmonic spring was monitored during the entire SMD simulation and plotted against the end-to-end distance. All MD simulations were performed using NAMD3 (76, 77). The system was described by the CHARMM36m force field (78, 79) for proteins and TIP3P for water (80). Long-range electrostatic forces were calculated using particle-mesh Ewald (PME) (81), and a distance cutoff of 12 Å was used to model the short-range, non-bonded interactions. The integration time was 2 fs, and the temperature was maintained at 310 K by Langevin thermostat with a damping coefficient of 1 ps^{-1}. All simulations were run as NPT ensembles.

Acknowledgements

We thank the Biochemistry Department Shared Instruments Pool at CU Boulder (RRID: SCR_018986) for use of its general infrastructure (centrifuges, Avestin C3 homogenizer, Applied Photophysics Chirascan Plus CD and stopped flow modules). We are immensely grateful for the support provided by SIP staff members Dr. Annette Erbse and Emily Proksch. This work was supported in part by NSF grant 2145848 to MCS and ET, NSF Grant MCB-2139572 to TTP, NIH grant R24GM145965 to ET and NIH Molecular Biophysics T32 Training Grant GM065103 supporting KED.

References:

1. D. M. Anderson, O. Schneewind, A mRNA signal for the type III secretion of Yop proteins by Yersinia enterocolitica. *Science* **278**, 1140-1143 (1997).
2. R. Arnold *et al.*, Sequence-Based Prediction of Type III Secreted Proteins. *PLoS Pathog* **5**, e1000376 (2009).
3. S. C. Birtalan, R. M. Phillips, P. Ghosh, Three-dimensional secretion signals in chaperone-effector complexes of bacterial pathogens. *Mol. Cell* **9**, 971-980 (2002).
4. Y. Luo *et al.*, Structural and biochemical characterization of the type III secretion chaperones CesT and SigE. *Nat Struct Biol* **8**, 1031-1036 (2001).
5. T. Michiels, P. Wattiau, R. Brasseur, J. M. Ruysschaert, G. Cornelis, Secretion of Yop proteins by Yersiniae. *Infect Immun* **58**, 2840-2849 (1990).
6. H. Russmann, T. Kubori, J. Sauer, J. E. Galan, Molecular and functional analysis of the type III secretion signal of the Salmonella enterica InvJ protein. *Mol Microbiol* **46**, 769-779 (2002).
7. Y. Akeda, J. E. Galán, Chaperone release and unfolding of substrates in type III secretion. *Nature* **437**, 911-915 (2005).
8. J. E. Galan, G. Waksman, Protein-Injection Machines in Bacteria. *Cell* **172**, 1306-1318 (2018).
9. P. C. Lee, A. Rietsch, Fueling type III secretion. *Trends Microbiol* **23**, 296-300 (2015).
10. A. Alonso *et al.*, Lck dephosphorylation at Tyr-394 and inhibition of T cell antigen receptor signaling by Yersinia phosphatase YopH. *J Biol Chem* **279**, 4922-4928 (2004).
11. D. S. Black, J. B. Bliska, Identification of p130Cas as a substrate of Yersinia YopH (Yop51), a bacterial protein tyrosine phosphatase that translocates into mammalian cells and targets focal adhesions. *EMBO J* **16**, 2730-2744 (1997).
12. S. Mühlen, M.-H. Ruchaud-Sparagano, B. Kenny, Proteasome-independent degradation of canonical NFkappaB complex components by the NleC protein of pathogenic Escherichia coli. *J. Biol. Chem.* **286**, 5100-5107 (2011).
13. H. Yen *et al.*, NleC, a Type III Secretion Protease, Compromises NF-κB Activation by Targeting p65/RelA. *PLoS Path.* **6**, e1001231 (2010).
14. J. S. Pearson, P. Riedmaier, O. Marches, G. Frankel, E. L. Hartland, A type III effector protease NleC from enteropathogenic Escherichia coli targets NF-κB for degradation. *Mol. Microbiol.* **80**, 219-230 (2011).
15. D. Tezcan-Merdol, L. Engstrand, M. Rhen, Salmonella enterica SpvB-mediated ADP-ribosylation as an activator for host cell actin degradation. *Int. J. Med. Microbiol.* **295**, 201-212 (2005).
16. Y. Fu, J. E. Galán, A salmonella protein antagonizes Rac-1 and Cdc42 to mediate host-cell recovery after bacterial invasion. *Nature* **401**, 293-297 (1999).
17. M. C. Schlumberger, W. D. Hardt, Triggered phagocytosis by Salmonella: bacterial molecular mimicry of RhoGTPase activation/deactivation. *Curr. Top. Microbiol. Immunol.* **291**, 29-42 (2005).
18. A. Peth, J. A. Nathan, A. L. Goldberg, The ATP costs and time required to degrade ubiquitinated proteins by the 26 S proteasome. *J Biol Chem* **288**, 29215-29222 (2013).

19. S. M. Doyle, O. Genest, S. Wickner, Protein rescue from aggregates by powerful molecular chaperone machines. *Nature Reviews Molecular Cell Biology* **14**, 617-629 (2013).
20. A. O. Olivares, H. C. Kotamarthi, B. J. Stein, R. T. Sauer, T. A. Baker, Effect of directional pulling on mechanical protein degradation by ATP-dependent proteolytic machines. *Proc Natl Acad Sci U S A* **114**, E6306-E6313 (2017).
21. A. R. Nager, T. A. Baker, R. T. Sauer, Stepwise unfolding of a beta barrel protein by the AAA+ ClpXP protease. *J Mol Biol* **413**, 4-16 (2011).
22. R. Zarivach, M. Vuckovic, W. Y. Deng, B. B. Finlay, N. C. J. Strynadka, Structural analysis of a prototypical ATPase from the type III secretion system. *Nat. Struct. Mol. Biol.* **14**, 131-137 (2007).
23. D. D. Majewski *et al.*, Cryo-EM structure of the homohexameric T3SS ATPase-central stalk complex reveals rotary ATPase-like asymmetry. *Nat Commun* **10**, 626 (2019).
24. V. T. Lee, O. Schneewind, Yop fusions to tightly folded protein domains and their effects on Yersinia enterocolitica type III secretion. *J Bacteriol* **184**, 3740-3745 (2002).
25. J. A. Sorg, N. C. Miller, M. M. Marketon, O. Schneewind, Rejection of impassable substrates by Yersinia type III secretion machines. *J Bacteriol* **187**, 7090-7102 (2005).
26. J. Radics, L. Konigsmaier, T. C. Marlovits, Structure of a pathogenic type 3 secretion system in action. *Nat Struct Mol Biol* **21**, 82-87 (2014).
27. K. Dohlich, A. B. Zumsteg, C. Goosmann, M. Kolbe, A substrate-fusion protein is trapped inside the Type III Secretion System channel in Shigella flexneri. *PLoS Pathog* **10**, e1003881 (2014).
28. J. E. Dawson, L. K. Nicholson, Folding kinetics and thermodynamics of Pseudomonas syringae effector protein AvrPto provide insight into translocation via the type III secretion system. *Protein Sci* **17**, 1109-1119 (2008).
29. M. A. LeBlanc, M. R. Fink, T. T. Perkins, M. C. Sousa, Type III secretion system effector proteins are mechanically labile. *Proc Natl Acad Sci U S A* **118** (2021).
30. K. Baruch *et al.*, Metalloprotease type III effectors that specifically cleave JNK and NF-kB. *EMBO J*, 1-11 (2010).
31. M. M. Turco, M. C. Sousa, The structure and specificity of the type III secretion system effector NleC suggest a DNA mimicry mechanism of substrate recognition. *Biochemistry* **53**, 5131-5139 (2014).
32. R. Arnold *et al.*, Sequence-based prediction of type III secreted proteins. *PLoS Pathog* **5**, e1000376 (2009).
33. N. D. Rawlings *et al.*, The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. *Nucleic Acids Res* **46**, D624-D632 (2018).
34. A. G. Murzin, S. E. Brenner, T. Hubbard, C. Chothia, SCOP: A structural classification of proteins database for the investigation of sequences and structures. *J Mol Biol* **247**, 536-540 (1995).
35. M. Punta *et al.*, The Pfam protein families database. *Nucleic Acids Res* **40**, D290-D301 (2011).
36. I. V. Demidyuk *et al.*, Crystal structure of the protealysin precursor: insights into propeptide function. *J Biol Chem* **285**, 2003-2013 (2010).
37. J. Yin, A. J. Lin, D. E. Golan, C. T. Walsh, Site-specific protein labeling by Sfp phosphopantetheinyl transferase. *Nat Protoc* **1**, 280-285 (2006).
38. M. Schlierf, F. Berkemeier, M. Rief, Direct observation of active protein folding using lock-in force spectroscopy. *Biophys J* **93**, 3989-3998 (2007).
39. C. Schoeler *et al.*, Mapping Mechanical Force Propagation through Biomolecular Complexes. *Nano Lett.* **15**, 7370-7376 (2015).
40. M. S. Bull, R. M. Sullan, H. Li, T. T. Perkins, Improved single molecule force spectroscopy using micromachined cantilevers. *ACS Nano* **8**, 4984-4995 (2014).
41. D. T. Edwards *et al.*, Optimizing 1-mu s-Resolution Single-Molecule Force Spectroscopy on a Commercial Atomic Force Microscope. *Nano Lett.* **15**, 7091-7098 (2015).
42. D. T. Edwards, T. T. Perkins, Optimizing force spectroscopy by modifying commercial cantilevers: Improved stability, precision, and temporal resolution. *Journal of Structural Biology* **197**, 13-25 (2017).
43. M. Rief, M. Gautel, F. Oesterhelt, J. M. Fernandez, H. E. Gaub, Reversible unfolding of individual titin immunoglobulin domains by AFM. *Science* **276**, 1109-1112 (1997).
44. F. Oesterhelt *et al.*, Unfolding pathways of individual bacteriorhodopsins. *Science* **288**, 143-146 (2000).
45. Y. Cao, H. Li, Polyprotein of GB1 is an ideal artificial elastomeric protein. *Nat Mater* **6**, 109-114 (2007).
46. E. Evans, K. Ritchie, Strength of a weak bond connecting flexible polymer chains. *Biophys J* **76**, 2439-2447 (1999).
47. E. Evans, K. Ritchie, Dynamic strength of molecular adhesion bonds. *Biophys. J.* **72**, 1541-1555 (1997).
48. H. Dietz, M. Rief, Exploring the energy landscape of GFP by single-molecule mechanical experiments. *P Natl Acad Sci USA* **101**, 16192-16197 (2004).
49. C. L. Chyan *et al.*, Reversible mechanical unfolding of single ubiquitin molecules. *Biophys J* **87**, 3995-4006 (2004).
50. S. R. Ainavarapu, L. Li, C. L. Badilla, J. M. Fernandez, Ligand binding modulates the mechanical stability of dihydrofolate reductase. *Biophys J* **89**, 3337-3344 (2005).
51. T. Hoffmann, K. M. Tych, M. L. Hughes, D. J. Brockwell, L. Dougan, Towards design principles for determining the mechanical stability of proteins. *Phys Chem Chem Phys* **15**, 15767-15780 (2013).
52. O. K. Dudko, G. Hummer, A. Szabo, Intrinsic rates and activation free energies from single-molecule pulling experiments. *Physical Review Letters* **96** (2006).
53. M. T. Woodside, S. M. Block, Reconstructing folding energy landscapes by single-molecule force spectroscopy. *Annu Rev Biophys* **43**, 19-39 (2014).

54. S. O. Garbuzynskiy, D. N. Ivankov, N. S. Bogatyreva, A. V. Finkelstein, Golden triangle for folding rates of globular proteins. *Proc Natl Acad Sci U S A* **110**, 147-150 (2013).
55. A. V. Glyakina, O. V. Galzitskaya, How Quickly Do Proteins Fold and Unfold, and What Structural Parameters Correlate with These Values? *Biomolecules* **10** (2020).
56. M. Sotomayor, K. Schulten, Single-molecule experiments in vitro and in silico. *Science* **316**, 1144-1148 (2007).
57. G. Arad-Haase *et al.*, Mechanical unfolding of acylphosphatase studied by single-molecule force spectroscopy and MD simulations. *Biophys J* **99**, 238-247 (2010).
58. M. Carrion-Vazquez *et al.*, The mechanical stability of ubiquitin is linkage dependent. *Nat Struct Biol* **10**, 738-743 (2003).
59. D. Sharma *et al.*, 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* **104**, 9278-9283 (2007).
60. J. Morfill *et al.*, Force-based analysis of multidimensional energy landscapes: application of dynamic force spectroscopy and steered molecular dynamics simulations to an antibody fragment-peptide complex. *J Mol Biol* **381**, 1253-1266 (2008).
61. W. Lee *et al.*, Full reconstruction of a vectorial protein folding pathway by atomic force microscopy and molecular dynamics simulations. *J Biol Chem* **285**, 38167-38172 (2010).
62. J. Schonfelder, R. Perez-Jimenez, V. Munoz, A simple two-state protein unfolds mechanically via multiple heterogeneous pathways at single-molecule resolution. *Nat Commun* **7**, 11777 (2016).
63. D. L. Guzman, A. Randall, P. Baldi, Z. Guan, Computational and single-molecule force studies of a macro domain protein reveal a key molecular determinant for mechanical stability. *Proc Natl Acad Sci U S A* **107**, 1989-1994 (2010).
64. T. Verdorfer *et al.*, Combining in Vitro and in Silico Single Molecule Force Spectroscopy to Characterize and Tune Cellulosomal Scaffoldin Mechanics. *J Am Chem Soc* 10.1021/jacs.7b07574 (2017).
65. S. P. Ng *et al.*, 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 (2005).
66. M. Iimori *et al.*, Structural basis of effector recognition by the T3SS chaperone VecA from Vibrio parahaemolyticus. *Biochem Biophys Res Commun* **776**, 152190 (2025).
67. M. Lilic, M. Vujanac, C. E. Stebbins, A Common Structural Motif in the Binding of Virulence Factors to Bacterial Secretion Chaperones. *Mol. Cell* **21**, 653-664 (2006).
68. F. D. Schubot *et al.*, Three-dimensional structure of a macromolecular assembly that regulates type III secretion in Yersinia pestis. *J Mol Biol* **346**, 1147-1161 (2005).
69. M. K. Hasan, S. El Qaidi, P. R. Hardwidge, The T3SS Effector Protease NleC Is Active within Citrobacter rodentium. *Pathogens* **10** (2021).
70. J. K. Faulk, D. T. Edwards, M. S. Bull, T. T. Perkins, Improved Force Spectroscopy Using Focused-Ion-Beam-Modified Cantilevers. *Methods Enzymol* **582**, 321-351 (2017).
71. R. Walder *et al.*, Rapid Characterization of a Mechanically Labile alpha-Helical Protein Enabled by Efficient Site-Specific Bioconjugation. *J Am Chem Soc* **139**, 9867-9875 (2017).
72. J. E. Sader, J. W. M. Chon, P. Mulvaney, Calibration of rectangular atomic force microscope cantilevers. *Rev Sci Instrum* **70**, 3967-3969 (1999).
73. B. Isralewitz, M. Gao, K. Schulten, Steered molecular dynamics and mechanical functions of proteins. *Curr Opin Struct Biol* **11**, 224-230 (2001).
74. S. Izrailev, S. Stepaniants, M. Balsera, Y. Oono, K. Schulten, Molecular dynamics study of unbinding of the avidin-biotin complex. *Biophys J* **72**, 1568-1581 (1997).
75. W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics. *J Mol Graph* **14**, 33-38, 27-38 (1996).
76. J. C. Phillips *et al.*, Scalable molecular dynamics with NAMD. *J Comput Chem* **26**, 1781-1802 (2005).
77. J. C. Phillips *et al.*, Scalable molecular dynamics on CPU and GPU architectures with NAMD. *J Chem Phys* **153**, 044130 (2020).
78. R. B. Best *et al.*, Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone phi, psi and side-chain chi(1) and chi(2) dihedral angles. *J Chem Theory Comput* **8**, 3257-3273 (2012).
79. J. Huang *et al.*, CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat Methods* **14**, 71-73 (2017).
80. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, Comparison of Simple Potential Functions for Simulating Liquid Water. *J Chem Phys* **79**, 926-935 (1983).
81. U. Essmann *et al.*, A Smooth Particle Mesh Ewald Method. *J Chem Phys* **103**, 8577-8593 (1995).