Mapping an Equilibrium Folding Intermediate of the Cytolytic Pore Toxin ClyA with Single-Molecule FRET

Fabian Dingfelder,[†] Stephan Benke,[†] Daniel Nettels,[†] and Benjamin Schuler*^{,†,‡}

[†]Department of Biochemistry, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland [‡]Department of Physics, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland

S Supporting Information

ABSTRACT: The 303-residue cytolytic toxin ClyA forms a stable α -helical monomer. In the presence of detergents or membranes, however, the protein makes a large conformational transition to the protomer state, which is competent for assembly into a dodecameric cytolytic pore. In this study, we map the structure of the ClyA monomer during denaturant-induced unfolding with single-molecule Förster resonance energy transfer (FRET) spectroscopy. To this end, we probe intramolecular distances of six different segments of ClyA by placing donor and acceptor fluorophores at corresponding positions along the chain. We identify an intermediate state that contains the folded core consisting of three of the α -helices that make up the helical bundle present in the structure of both the monomer and the protomer, but with the C- and N-terminal helices unfolded, in accord with the secondary structure content estimated from circular dichroism (CD) spectroscopy. The existence of this intermediate is likely to be a consequence of the



structural bistability underlying the biological function of ClyA: The terminal helices are part of the largest rearrangements during protomer formation, and the local differences in stability we detect may prime the protein for the required conformational transition.

■ INTRODUCTION

During the past two decades, our understanding of the elementary processes of protein folding has improved tremendously. These advances are largely due to a focus on single-domain proteins that typically exhibit fast folding and can be well described by two-state models that enable a quantitative thermodynamic and kinetic analysis.^{1,2} By dissecting such systems into their secondary structure building blocks, combined with methods from ultrafast spectroscopy, other experimental techniques, and the continued advances in simulation and theory,³⁻⁶ we now have a remarkably detailed picture of many protein folding processes from picoseconds to milliseconds. These developments in conceptual understanding and experimental tools now enable us to revisit questions that had been the focus of protein folding research in its early days, where most of the model systems were relatively large, multidomain, and often oligomeric proteins.⁷ Already 40 years ago, it was clear that many of these proteins populate folding intermediates that could be detected experimentally,⁸ and many of the ideas and methods used to investigate biomolecules to this day were established for identifying and characterizing such intermediates. However, we can now take advantage of methods such as single-molecule spectroscopy,^{9–16} which were not available at that time.

Folding intermediates have been observed under a variety of conditions, such as low pH or at moderate concentrations of denaturant.^{7,8,17,18} Alternatively, equilibrium intermediate states can be populated by introducing mutations that destabilize the native state or by the removal of ligands. For

example, CD spectroscopy of cytochrome c, β -lactamase, and apomyoglobin revealed an acid-denatured state around pH 2, which folds to an intermediate state upon further reduction in pH. The characteristics of this compact state were a high content of secondary structure (as identified by far-UV CD) and few tertiary interactions (near-UV CD), a constellation that is often referred to as a molten globules state.^{18,19} In this and many other cases, however, obtaining more detailed structural information on intermediate states from ensemble techniques is often challenging, especially under conditions where the native state, the intermediate, and the unfolded state coexist. A technique that allows structural information on intermediates to be obtained even at atomic resolution is nuclear magnetic resonance (NMR) spectroscopy. For example, the structure of the intermediate of a small fourhelix bundle protein that is sparsely populated at equilibrium (2-3%) could be solved by NMR.²⁰ However, obtaining highresolution information from NMR becomes increasingly challenging for larger proteins. Another method that enables folding intermediates to be mapped is single-molecule fluorescence spectroscopy, especially in combination with FRET.^{9,13,21,22} A particular benefit of this technique is the ability to resolve individual subpopulations, so that the

Special Issue: William A. Eaton Festschrift

 Received:
 July 22, 2018

 Revised:
 August 28, 2018

 Published:
 August 29, 2018

intermediate populations can be separated from the unfolded or native state. Furthermore, single-molecule FRET is a versatile approach for the investigation of protein kinetics and dynamics²³ and the properties of conformationally heterogeneous systems, such as unfolded and intrinsically disordered proteins²⁴ or misfolded states.^{25,26} Even though the focus of most single-molecule FRET experiments has been the folding of small proteins, the approach has clearly been shown to have the potential for resolving structural and kinetic heterogeneity in the folding of multidomain and oligomeric proteins.^{25–30}

In this study, we map the structure of an equilibrium intermediate of the large protein cytolysin A (ClvA, 34 kDa) with single-molecule FRET spectroscopy by placing donoracceptor pairs into different positions to monitor distance changes in different parts of the protein. ClyA is a poreforming bacterial toxin that is expressed as a highly α -helical monomer. In the presence of membranes or detergents, the protein undergoes a very large structural transformation into the protomer conformation (Figure 1), which then assembles into the cytolytic pore.^{30–38} The ClyA monomer is a bundle of four long helices (αA , αB , αC , and αF) flanked by a shorter Cterminal helix $(\alpha G)^{39}$ Membrane binding is thought to be triggered by its interaction with the " β -tongue", a short hydrophobic β -hairpin at one end of the protein (Figure 1a).³² Upon protomer formation, the amphipathic N-terminal helix α A largely dislocates from the helix bundle of the monomer and flips upward by 180° to become an extension of helix αB .³² Concomitantly, large parts of the hydrophobic core are repacked, and the arrangement and extent of the other helices change such that more than half of all residues are affected by the reorganization of the molecule. ClyA thus represents a remarkable bistable system that can interconvert between two very different conformations. The kinetic mechanism of the monomer-to-protomer conversion has previously been investigated, which revealed that the process involves an expanded intermediate state.³⁰ As for other such metamorphic proteins,² the existence of two stable conformational states for the same polypeptide sequence suggests a delicate balance of interactions,^{28,41} which for a large protein such as ClyA may be expected to lead to detectable differences in conformational stability across the protein. Here we thus investigate ClyA for the presence of equilibrium intermediates that may be related to the conformational change essential for its biological function.

METHODS

Protein Expression, Purification, and Labeling. ClyA was expressed, purified, and labeled as described previously.³⁰ Briefly, six double-cysteine variants (A2C/Q56C, K107C/N157C, A187C/K290C, Q56C/E252C, T2C/E252C, and T2C/V303C) were produced by site-directed mutagenesis of the pClyA vector coding for N-terminally His₆-tagged ClyA,³³ using the QuickChange protocol (Stratagene). Protein expression was carried out in *E. coli* Tuner DE3 (Merck Millipore) at 20 °C for 12 h. The double-cysteine variants as well as wt-ClyA were purified via immobilized metal affinity chromatography (IMAC) with a nickel-nitrilotriacetic acid (NTA) resin (Thermo Fischer).

wt-ClyA for the CD experiments was buffer-exchanged to 10 mM sodium phosphate (pH 7.4) on a HiPrep desalting 26/100 column (GE Healthcare) and further purified by anion exchange chromatography on a HighPrep Q FF 16/100 (GE Healthcare) applying an elution gradient of 0-1 M NaCl.³⁰



Figure 1. Crystal structures of (a) the ClyA monomer (PDB code: $1QOY^{39}$) and (b) the protomer conformation (PDB code: $2WCD^{32}$). The helices from αA to αG are indicated in different colors consistent between the two structures. Dashed lines indicate segments that were not resolved in the crystal structures and are thus likely to be disordered. These are residues 299–303 in the monomer and residues 1–7 as well as 293–303 in the protomer. (c) Contact map (cutoff distance 6.5 Å) of the monomer (magenta) and the protomer (green). Contacts present in both conformations are shown in black. The maps were generated with CMView.⁵⁷ Inset: Structure of the ClyA monomer. The coloring scheme (see scale) reflects the displacement of residues that are resolved in both crystal structures. The residues indicated in black (a-c) were exchanged to cysteines for fluorescence labeling. Protein structures were generated with Pymol.⁵⁸

The protein was then denatured by adding guanidinium chloride (GdmCl) to a final concentration of ~6 M and concentrated to 230 μ M protein. For final purification, size exclusion chromatography on a Superdex 200 Increase 10/300 column (GE Healthcare) was performed, yielding a final ClyA concentration of 50 μ M.

Following nickel chelate affinity chromatography, the double-cysteine variants for labeling were concentrated to ~400 μ M and reduced by adding Dithiothreitol (DTT) to a final concentration of 10 mM. The reducing agent was removed by a HiTrap desalting column (GE Healthcare) using a buffer containing 50 mM sodium phosphate and 150 mM NaCl (pH 7.3). For donor labeling, Alexa Fluor 488 C5 maleimide (Invitrogen) dissolved in dimethyl sulfoxide (DMSO) was added in a dye-to-protein molar ratio of 1:0.7 with the protein concentration being ~100-300 μ M. The reaction was carried out at room temperature for 2 h and then quenched by the addition of DDT to a final concentration of 50 mM. After removal of free dye and buffer exchange to

10 mM sodium phosphate (pH 7.3) via a HiTrap desalting column (GE Healthcare), the singly labeled protein was separated from unlabeled and doubly labeled protein by anion exchange chromatography, using a MonoQ GL anion exchange column (GE Healthcare). The protein was concentrated to ~100 μ M, and Alexa Fluor 594 C5 maleimide (Invitrogen), dissolved in DMSO, was added in 3-fold molar excess. The labeling reaction was carried out overnight on ice and quenched by the addition of DDT to a final concentration of 50 mM. Unreacted dye was removed by a HiTrap desalting column, followed by anion-exchange chromatography as in the first labeling step. The masses of all doubly labeled variants were confirmed by electrospray ionization mass spectrometry.

CD Spectroscopy. Equilibrium spectra (Figure 3a) were recorded on a Jasco J715 CD spectrometer at 293 K. The 50 μ M ClyA stock solution in 4 M GdmCl was diluted 11-fold with buffer (25 mM sodium phosphate, 75 mM NaCl, pH 7.0) containing different amounts of GdmCl. The samples were incubated for at least 1 h, and data were recorded in a 1 mm path length cuvette in the range of 200–250 nm. The scanning speed was 10 nm/min, and 15 spectra were averaged. For the measurement in 4 M GdmCl, a 10- μ M sample was measured in a 0.5 mm path length cuvette, and 20 accumulations were averaged.

Singular Value Decomposition (SVD). SVD was used to determine the number of components necessary to account for the changes in the CD spectra as a function of denaturant concentration.⁴² CD spectra measured at N different denaturant concentrations were combined into an $M \times N$ matrix \mathbf{H} , where M is the number of wavelengths. This matrix can be decomposed into the product of three matrices: H =**USV**^T. The orthogonal $M \times M$ matrix **U** contains information about the shape of the CD spectra, with the columns of U being the basis vectors (Figure 3c). The diagonal matrix S contains the singular values sorted by magnitude as diagonal elements (inset Figure 3c). The columns of the orthonormal matrix V contain information regarding the change in signal amplitude as a function of denaturant concentration (amplitude vectors). The first two amplitude vectors were fitted globally with a two-state model to obtain the unfolding midpoint denaturant concentration, with the midpoint as a shared fit parameter (Figure 3d).

Single-Molecule Measurements. Single-molecule measurements were performed on a Micro Time 200 confocal instrument (PicoQuant) at 295 K, as described previously.⁴² In brief, pulsed interleaved excitation⁴³ (PIE) was applied to remove the contribution of molecules lacking an active acceptor dye. The fluorescence photons emitted after donor or acceptor excitation were collected via an UplanApo 60/ 1.2W objective (Olympus) and distributed onto four singlephoton avalanche detectors (τ -SPAD, PicoQuant), first according to polarization (polarization cube, PicoQuant) and then according to color (595 DCXR dichroic mirror, Chroma). Prior to detection, photons were filtered by E252/50 M or HQ650/100 band-pass filters (Chroma) for donor and acceptor channels, respectively. The protein stock solution of the different ClyA variants in 4 M GdmCl was diluted at least 100-fold with buffer (25 mM sodium phosphate, 75 mM NaCl, 140 mM β -mercaptoethanol, 0.001% (w/v) Tween 20, pH 7.0) containing different amounts of GdmCl, resulting in sample concentrations of ~40-125 pM. The Tween 20 concentration used to reduce surface adhesion is below the critical micelle concentration $(5.9 \times 10^{-5} \text{ M}; 0.007\% (w/v))^{44}$

and does not induce the conversion to the protomer conformation.³⁰ The samples were incubated for at least 2 h for equilibration prior to measurements. Measurements were conducted in custom-made quartz cuvettes that were plasma-activated (FEMTO, Diener) and pretreated at least for 15 min with 0.1 mg/mL poly(L-lysine)-graft-poly(ethylene glycol) (PLL/PEG) (PLL(20)-g[3.5]-PEG(2), Susos) to minimize surface adhesion of ClyA molecules. Measurements were carried out for 30 min (Figure 4).

Single-Molecule FRET Data Analysis. FRET data analysis was carried out as described previously.⁴² The numbers of detected donor and acceptor photons were corrected for the different quantum yields of the dyes, different detection efficiencies, crosstalk, acceptor direct excitation, and background.⁴⁵ Contiguous photons detected after donor excitation pulses with interphoton times of less than 100 μ s were combined into one fluorescence burst. Bursts were considered for data analysis if more than 50 photons were identified. Bursts with a stoichiometry ratio S > 0.7 were omitted. For each burst, the transfer efficiency was calculated according to $E = n_A/(n_A+n_D)$, with n_A and n_D being the numbers of corrected acceptor and donor photons detected after donor efficiencies were binned in a histogram.

To extract the relative fractions of the native, intermediate, and unfolded populations for each variant, the histograms were fitted with a sum of appropriate peak functions, one for each subpopulation.⁴² The width of each subpopulation peak was a shared fit parameter throughout each titration series, while the peak positions and amplitudes were allowed to be different in each in histogram (Figure 4). The uncertainty in the peak populations was estimated by increasing and decreasing the width of one of the populations by 15% for the fit of the transfer efficiency histograms, with the widths of the other populations remaining unchanged.⁴² This procedure was iteratively applied to all populations contributing to the transfer efficiency histogram, and the lowest and highest values of the resulting populations are reported as error bars (Figure 5).

Dual-Focus Fluorescence Correlation Spectroscopy (2f-FCS). 2f-FCS measurements were performed on a confocal single-molecule instrument (MT200, PicoQuant), with two foci (interfocal distance of 469 nm, calibrated as previously described⁴⁶) produced by passing two collinear, alternatingly pulsed laser beams of orthogonal polarization through a differential interference contrast prism⁴⁷ (U-DICTHC, Olympus). The light sources were a SC-450-4 supercontinuum laser (Fianium, UK) with a z582/15 band-pass filter, resulting in a wavelength of 582 \pm 15 nm, and a Solea supercontinuum laser (PicoQuant, Germany) with wavelength-selected output at 585 \pm 5 nm. The repetition rate of each laser was 20 MHz, and the laser power was 30 μ W (measured at the back aperture of the objective). Fluorescence emission, collected by the objective, was focused on a 150- μ m pinhole before being distributed by a 50/50 beam splitter onto two single-photon avalanche photodiodes (τ -SPAD 50, PicoQuant) after passing a bandpass filter (ET525/50M, Chroma). The photon arrival times were recorded by two channels of a HydraHarp 400 counting module (PicoQuant). Fluorescence intensity auto- and crosscorrelation curves from the data of each focus were fitted globally as described by Pieper et al.⁴⁷ From the resulting translational diffusion coefficient, D, the Stokes radius was calculated according to the Stokes-Einstein equation: $r_{\rm S}$ =

 $k_{\rm B}T/6\pi\eta D$, with $k_{\rm B}$ being the Boltzmann constant, *T* being the temperature, and η being the solvent viscosity. Here we accounted for the dependency of η on the GdmCl-concentration of the sample (Figure S2).⁴⁶

Nanosecond Fluorescence Correlation Spectroscopy (nsFCS). Autocorrelation curves of donor and acceptor channels as well as cross-correlation curves between the donor and acceptor channels were measured as described previously.⁴⁸ The three curves $g_{AA}(\tau)$, $g_{DD}(\tau)$, and $g_{DA}(\tau)$ were globally fitted in the range of lag times $\tau = -100$ ns to $\tau = 100$ ns with

$$g_{ij}(\tau) = 1 + \frac{1}{N} (1 - c_{ab} e^{-|\tau|/\tau_{ab}}) (1 + c_{cd} e^{-|\tau|/\tau_{cd}}), \quad i, j = A, D$$
(1)

where *N* is the mean number of molecules in the confocal volume. The two exponentially decaying components represent photon antibunching (ab) and conformational dynamics (cd), respectively. The corresponding relaxation times τ_{ab} and τ_{cd} were shared fit parameters of the global fit, whereas *N* and the amplitudes, c_{ab} and c_{cd} , were fitted individually for each correlation curve.⁴⁹ τ_{cd} was converted to the reconfiguration time (i.e., the relaxation time of the interdye distance correlation function) as previously described⁴⁹ using the distance distribution of a self-avoiding-walk chain⁵⁰ and a Förster radius of 5.6 nm (Figure 6).

Calculation of Mean Transfer Efficiencies from Interdye-Distance Dynamics in a Potential of Mean Force. The spatial distribution of sterically accessible dye positions, taking into account the flexible linkers, was modeled using the software FRET Positioning and Screening (FPS)⁵¹ (Figure 2). As input, the structure of the ClyA monomer (PDB code: 1QOY) and the dimensions of the fluorescent dyes (Alexa Fluor 488 C5 maleimide and Alexa Fluor 594 C5 maleimide) and the linker lengths and widths are needed. The dye dimensions of Alexa Fluor 488 and Alexa 594 were estimated to be 5, 4.5, and 1.4 Å as well as 8.1, 4.2. and 2.1 Å, respectively, from their chemical structures. The linker lengths and widths used were 16 and 4.5 Å (Alexa Fluor 488 C5 maleimide) as well as 18 and 4.5 Å (Alexa Fluor 594 C5 maleimide). From the output, an expected interdye-distance equilibrium distribution $P_{eq}(r)$ can be derived. For calculating the corresponding mean transfer efficiency, we took into account rapid fluctuations of the interdye distance that are expected to occur on time scales similar to the fluorescence lifetime of the donor (\sim 4 ns in the absence of the acceptor). We modeled the dynamics as diffusive motion in the potential of mean force given by the distance distribution. Following the theory and notation of Gopich and Szabo,⁵² we briefly describe the procedure: Very generally, the following relation between the mean transfer efficiency $\langle E \rangle$ and the mean lifetime $\langle \tau^* \rangle$ of the excited state of the donor dye holds

$$\langle E \rangle = 1 - \langle \tau^* \rangle / \tau_D \tag{2}$$

where $\tau_D = 1/k_D$ is the mean fluorescence lifetime of the donor in the absence of an acceptor. $\langle \tau^* \rangle$ is given by

$$\langle \tau^* \rangle = \int_0^\infty p_{D^*}(t) \mathrm{d}t \tag{3}$$

Here, $p_{D^*}(t)$ is the survival probability of the excited state of the donor (D^*), which formally can be written as⁵²

$$p_{D^*}(t) = \left\langle \exp\left(-k_D t - \int_0^t k_D \left(\frac{R_0}{r(t')}\right)^6 dt'\right) \right\rangle_{r(t)}$$
(4)

where $\langle ... \rangle_{r(t)}$ denotes averaging over all possible paths of distance fluctuations, r(t). A related quantity is the population density, $p_{D^*}(r,t)$, of D^* , which is given by the solution of the Smoluchowski equation

$$\frac{\partial p_{D^*}(r,t)}{\partial t} = \left[\mathcal{L} - k_D - k_D \left(\frac{R_0}{r}\right)^6 \right] p_{D^*}(r,t)$$
(5)

with the diffusion operator

$$\mathcal{L} = D \frac{\partial}{\partial r} P_{eq}(r) \frac{\partial}{\partial r} (P_{eq}(r))^{-1}$$
(6)

where *D* is the effective diffusion coefficient. The Smoluchowski equation needs to be solved with the initial condition $p_{D^*}(r,t=0) = P_{eq}(r)$, assuming that at t=0 the donor was excited and that the interphoton time is much longer than the reconfiguration time of the interdye distance. From the solution we obtain the survival probability after integration over *r*:

$$p_{D^*}(t) = \int_0^\infty p_{D^*}(r, t) dr$$
(7)

For the calculations in practice, we discretized the Smoluchowski equation with respect to r on a grid of N = 50 distances defined on an interval (r_a, r_b) outside of which $P_{eq}(r)$ is zero to good approximation. We obtain i = 1, ..., N rate equations of the form

$$\frac{\mathrm{d}p_i}{\mathrm{d}t} = \frac{D}{\Delta^2} \left(\frac{P_{i-1/2}}{P_{i-1}} p_{i-1} - \frac{P_{i-1/2} + P_{i+1/2}}{P_i} p_i + \frac{P_{i+1/2}}{P_{i+1}} p_{i+1} \right) - k_D (1 + (R_0/r_i)^6) p_i$$
(8)

where $r_i = r_a + (i-1)\Delta$ with $\Delta = (r_b - r_a)/(N-1)$. We further defined $p_i = p_{D^*}(r_i, t)$ and $P_i = P_{eq}(r_i)$. In the last formula, *i* is not limited to integers. At the boundaries, we set $p_0 = p_{N+1} =$ $P_{1/2} = P_{N+1/2} = 0$. The effective diffusion coefficient D is the sum of the diffusion coefficients of the individual dyes including their linkers attached to the protein. These values are difficult to measure directly. Here we assume that D is similar to the diffusion coefficient of a free dye, which is a reasonable approximation based on molecular dynamics simulations.⁵³ The experimentally determined diffusion coefficient of Alexa Fluor 488 was used (435 μ m²/s).⁵⁴ To estimate error bounds (Figure 2c), D was varied from a value corresponding to distance relaxation much slower than the fluorescence lifetime (static regime) to twice the value of the diffusion coefficient of the free dye, which corresponds to the largest physically meaningful value. To account for the two possible labeling permutations, the distance distributions and mean transfer efficiencies were calculated for both permutations and were found to be in good agreement.

To test the free rotational mobility of the fluorescent dyes, we determined fluorescence polarization anisotropies for the six different variants from the single-molecule data and found values between 0.12 and 0.14 for variants 2/56, 187/290, 56/252, 2/252, and 2/303, suggesting rapid orientational averaging of the dyes and thus an orientational factor of $\kappa^2 \approx 2/3$. For variant 107/157, we observed a donor anisotropy

of 0.25. In this case, reduced orientational averaging may contribute to the discrepancy between experimentally observed values and those calculated based on the accessible volumes. The donor fluorescence anisotropies decrease with increasing denaturant concentration, with values between 0.03 to 0.06 in 4 M GdmCl. Only for variant 107/157, we still observe an increased anisotropy of 0.16. As expected, the assumption of free and rapid orientational averaging is thus even more likely to hold under denaturing conditions.

Equilibrium Three-State Model. The three-state model used for fitting the relative populations of the native (N), intermediate (I), and unfolded state (U) (Figure 5) is given by K_1 , K_2

 $K_1
equiv K_2$ $N \rightleftharpoons I \rightleftharpoons U$ with the equilibrium constants $K_1 = c_I/c_N$ and $K_2 = c_U/c_I$. The fractions of the populations are then described by $f_N = (1 + K_1 + K_1K_2)^{-1}$, $f_I = (K_1^{-1} + 1 + K_2)^{-1}$, and $f_U = (K_1^{-1}K_2^{-1} + K_2^{-1} + 1)^{-1}$. The equilibrium constants depend on denaturant concentration c_{GdmCl} according to $K_{i=1,2} = \exp(-m_i(c_{\text{GdmCl}} - c_{m,i})/RT)$, with R being the ideal gas constant, T being the temperature, $c_{m,i}$ being the midpoint concentrations, and m_i being the equilibrium m-values.^{2,55} Fit parameters were the m-values and the midpoint concentrations.

RESULTS

Mapping the Structure of ClyA with Single-Molecule FRET. For mapping the structural changes during unfolding of ClyA, six fluorescently labeled variants of the 303-residue protein were designed for single-molecule FRET based on the crystal structure of the soluble ClyA monomer (PDB code: 1QOY³⁹) and the protomer (taken as one subunit of the pore, PDB code: 2WCD³²) (Figure 1a and b). By placing the FRET dyes into different positions, different parts of the structure with very different segment lengths (ranging from 50 to 301 residues) were probed (Figure 2). To minimize effects on conformational stability, the residues exchanged to cysteines for fluorescence labeling were chosen in a way that they are solvent-exposed both in the monomer and protomer conformations (Figure 1).

First, we recorded transfer efficiency histograms with confocal single-molecule detection of freely diffusing sample molecules for all ClvA variants under native conditions. The measured mean transfer efficiencies $\langle E \rangle$ were compared to the values calculated based on the crystal structure (Figure 2). The accessible volumes of the dyes, considering their flexible linkers, were modeled as described by Seidel and co-workers^{51,56} and are depicted in blue and orange for Alexa Fluor 488 (donor) and Alexa Fluor 594 (acceptor), respectively in Figure 2b. The measured and calculated values of $\langle E \rangle$ for the variants are compared in Figure 2c. The calculations were performed assuming different time scales for interdye-distance dynamics with respect to the fluorescence lifetime of the donor. We note that the dynamics have a detectable effect on the resulting transfer efficiencies,⁵⁹ especially if the distance between the attachment points of the linkers is similar to or smaller than the linker lengths (see Methods). Donor fluorescence anisotropies were determined for all variants to test for free rotational mobility of the dyes. Only for variant 107/157, we find an anisotropy in excess of 0.2, potentially contributing to the discrepancy between the measured and calculated mean transfer efficiency (see Methods). The last five C-terminal residues of ClyA are not resolved in the crystal structure of the monomer and are



Figure 2. Mapping of the native state of ClyA with single molecule FRET. (a) Transfer efficiency histograms of the six different variants investigated, fitted with appropriate peak functions (lines) to obtain mean transfer efficiencies, $\langle E_{meas} \rangle$. (b) Crystal structure of ClyA (PDB code: $100Y^{39}$) with the accessible volumes⁵¹ of the fluorophores depicted in blue and orange, for the donor and acceptor dyes, respectively. The part of the polypeptide chain that is probed in each variant is colored in magenta. (c) Comparison of measured and calculated mean transfer efficiency values based on the structure. Interdye-distance dynamics are modeled as diffusive process in a potential of mean force (see text). The error range for $\langle E_{calc} \rangle$ results from the limiting cases of fast diffusion (upper bound) or no diffusion of the dyes (lower bound) with respect to the fluorescence lifetime (see Methods for details). There might be additional contributions to the error on $\langle E_{calc} \rangle$, originating, e.g., from residual interactions of the dyes with the protein surface. However, those contributions are difficult to quantify and are thus not included here. Note that the five C-terminal residues are not resolved in the crystal structure, so the length of the flexible dye linker was increased for calculating the expected mean transfer efficiency for variant 2/303 (*).

presumably unstructured. Therefore, for the calculation of variant A2C/V303C, we added the length of five times 0.38 nm to the contour length of the C-terminal dye linker (bottom panel in Figure 2b), which leads to an increased uncertainty in the $\langle E \rangle$ estimate. The larger width of the transfer efficiency peak of this variant compared to the others further indicates an additional distance heterogeneity persistent on a millisecond time scale. Altogether, the distance values from the measurements and those based on the crystal structure are in reasonable agreement and illustrate the level of accuracy



Figure 3. Equilibrium folding experiments of wt-ClyA with CD spectroscopy. (a) Molar mean residue ellipticity as a function of wavelength plotted for different denaturant concentrations (see color scale). Note that at elevated GdmCl concentrations, increasing absorption of the denaturant prevented measurements at wavelengths lower than 210 nm. (b) Relative helical content as a function of denaturant concentration determined from the ellipticity values recorded at 222 nm. The error on the midpoint concentration (c_m) corresponds to the standard error of the fit. (c and d) SVD analysis of the spectra shown in (a). The first four basis vectors identified by SVD are shown in (c) together with the singular values of all components (inset). We regard only the first three as significant. (d) Amplitude dependence of the first four components on denaturant concentration. The first two components (left column) are globally fitted with an equilibrium two-state model to obtain c_m. The standard error of the fit is reported.

attainable in the single-molecule FRET experiments of this system.

Secondary Structure Changes from CD Spectroscopy. To quantify the secondary structure content of ClyA as a function of denaturant concentration, equilibrium CD spectra of unlabeled wt-ClyA were recorded (Figure 3a). Starting from 50 µM ClyA in 4 M GdmCl, the protein was diluted 11-fold with buffers containing various concentrations of the denaturant GdmCl and equilibrated. From the CD spectrum recorded at the lowest GdmCl concentration (0.36 M), where the protein is folded, the secondary structure content was quantified using the K2d algorithm⁶⁰ (provided by the Web server DICHROWEB), which is based on a reference database of spectra of proteins with known crystal structure.^{61,62} The estimated helicity was 81%, within error in agreement with a helical content of 77% determined from the crystal structure. Above ~4 M GdmCl, all helical signal is lost, indicating complete unfolding.

To analyze the unfolding transition, we first plotted the normalized mean-residue molar ellipticity at 222 nm as a function of denaturant concentration (Figure 3b). The resulting data show a single transition with surprisingly low cooperativity for a protein of this size;⁶³ fitting the transition with an equilibrium two-state model results in a midpoint concentration of 2.29 \pm 0.03 M GdmCl. The ellipticity at 222 nm does not reveal direct evidence for the existence of an equilibrium intermediate. Hence, to use the information from the entire spectral range, we quantify the number of components required to describe the data set using SVD.^{42,64} At least three singular values above the noise were identified

(Figure 3c), indicative of the existence of more than two states. However, the third and fourth basis vectors contain substantially higher levels of noise than the first two (Figure 3c), and their amplitude vectors are dominated by noise (Figure 3d), whereas the amplitude vectors of the first two components resemble the transition obtained from the single-wavelength analysis (Figure 3b). Fitting the first two amplitude vectors globally with a two-state model, using the midpoint concentration as a shared fit parameter, yielded an unfolding midpoint of 2.25 ± 0.10 M GdmCl, in agreement with the single-wavelength analysis. Overall, the results obtained from CD spectroscopy provide hints for a deviation from simple two-state behavior but remain rather inconclusive regarding the properties of a potential unfolding intermediate.

Mapping the Intermediate State of ClyA with Single-Molecule FRET. To test the existence of an intermediate in more detail, transfer efficiency histograms of the six FRETlabeled ClyA variants were recorded at various denaturant concentrations (Figure 4). The protein stock solutions in 4 M GdmCl were diluted at least 100-fold with the respective GdmCl buffer. The samples were then incubated for at least 2 h to equilibrate. (The shape of transfer efficiency histograms did not change over the time course of a measurement (Figure S1), confirming completed equilibration). In all cases, the transition proceeds from the folded state under fully native conditions to a single unfolded population at 4 M GdmCl, in agreement with the loss of CD signal at this denaturant concentration (Figure 3b). However, in contrast to the CD results, the existence of an intermediate state becomes obvious in the single-molecule experiments by the appearance of an additional population that is most prominent in the range between about 1 and 2.5 M GdmCl (Figure 4), especially for variants 187/290, 56/252/, 2/252, and 2/303. Interestingly, the patterns of transfer efficiency peaks are different for the different variants, both in terms of the number of peaks and the denaturant concentrations where they are populated. What can we infer about the structure of the intermediate from the single-molecule FRET measurements?

ClyA variant 2/56 (Figure 4a), which probes the N-terminal helix αA , shows a pronounced transition from the low transfer efficiency expected for the folded helix (Figure 2, first row) to a population with $\langle E \rangle \approx 0.4$ at a GdmCl concentration of \sim 1.25 M. No clearly separated third peak is observed at \sim 2.5 M, but some peak broadening is detected, indicating the existence of two overlapping subpopulations with similar transfer efficiencies. The mean interdye distance of the intermediate is thus close to that of the unfolded state for these labeling positions, suggesting that the N-terminal helix αA is unfolded in the intermediate. The small shift in $\langle E \rangle$ at ~2.5 M GdmCl suggests that part of the segment 2-56 remains structured in the intermediate state (possibly involving the second helical segment of αA , $\alpha A2$, after Pro36 and the following short 3_{10} helix, see Figure 1a) and only unfolds in the second transition along with the rest of the protein.

In contrast, ClyA 107/157, which probes helix α C, reveals a very different behavior. For this variant, an intermediate population is not detectable, and apparent two-state behavior is observed, with a single transition at ~2.25 M GdmCl (Figure 4b), implying that helix α C remains folded in the intermediate. This hypothesis is supported by the similar behavior of ClyA 56/252 (Figure 4d), which probes the segment encompassing helices α B- α F, i.e., the core of the protein without the N- and C-terminal helices. In this case, the mean transfer efficiency in



Figure 4. Mapping of the intermediate state of ClyA with single molecule FRET. (a-f) Transfer efficiency histograms of the six ClyA variants measured at increasing concentrations of denaturant (as indicated). Native, intermediate, and unfolded states subpopulations were fitted with suitable peak functions⁴² to obtain relative populations (Figure 5) and mean transfer efficiencies. The gradual shift of the unfolded population and the intermediate state population (if parts of the polypeptide chain are probed that contain unfolded segments) toward lower transfer efficiency with increasing denaturant concentration precluded a global fit of each titration series with the positions and widths of the populations as shared fit parameters. Therefore, only the width of each population was constrained to a constant value throughout each titration series, while peak positions were allowed to vary between histograms. Dashed lines indicate the mean transfer efficiencies of the populations assigned to the intermediate.



Figure 5. (a-f) Relative populations of the folded, unfolded, and intermediate states of the six variants as a function of denaturant concentration. Error bars reflect the uncertainty in fitting the transfer efficiency histograms (see Methods). For each variant an equilibrium three-state model (two-state model in b) was fitted to the data (lines). The cartoons indicate the probed segment of the polypeptide chain in magenta for each variant.

11257

the intermediate is only slightly increased compared to the native state, suggesting that helices α B- α F remain largely structured. Only above ~2 M GdmCl do we observe the transition to the low transfer efficiency population expected for

a fully unfolded sequence with a dye separation of 200 residues.

Variant 187/290 probes the segment including helices α F and α G. Interestingly, in this case the transfer efficiency of the intermediate is lower than for the fully unfolded state,



Figure 6. Probing nanosecond dynamics of the intermediate state with ns-FCS. ClyA 2/56 (a) and 56/252 (b) were measured at 1.75 M GdmCl. Depicted are the autocorrelations of acceptor and donor fluorescence (red and green, respectively) and the cross-correlation (orange). All correlation curves show the usual sharp decrease at short lag times due to photon antibunching. An additional component on longer time scales that is correlated in the autocorrelations and anticorrelated in the cross-correlation is indicative for interdye distance dynamics. This signature is clearly visible for variant 2/56, where helix α A is probed. A global fit (black lines) reveals a time constant of 31 ± 2 ns. Variant 56/252, probing the core helices, shows no clear indication for interdye distance dynamics on the tens of nanoseconds time scale.

indicating a very large interdye distance compared to both the native and the unfolded state (Figure 4c). Since we already know from variant 56/252 that helix α F remains largely intact in the intermediate state, we conclude that only the C-terminal helix α G unfolds in the first transition. The combination of an extended and still stably folded helix α F with an unfolded helix α G would explain the large distance between residues 187 and 290, which decreases upon unfolding of helix α F in the second transition to the fully unfolded state.

With variant 2/252, the full polypeptide chain with the exception of the C-terminal helix αG is probed (Figure 4e). In the first transition, the intermediate population appears at $\langle E \rangle \approx 0.5$. This value is in accord with helix αA being largely unfolded in the intermediate (as suggested based on variant 2/56) and the core helices $\alpha B \cdot \alpha F$ retaining their structure and arrangement (as suggested by variant 56/252). Finally, in the terminally labeled variant 2/303 (Figure 4f), the intermediate appears at an efficiency of ~0.67, presumably corresponding to helices αG and $\alpha A1$ being unfolded. The mean transfer efficiency of the native state population shifts continuously toward lower transfer efficiency upon increasing denaturant concentrations, which may indicate noncooperative unfolding or denaturant-induced expansion of the C-terminal 12-residue segment that follows helix αG .

To quantify the relative populations of the native, intermediate, and unfolded states as a function of denaturant concentration, the transfer efficiency histograms were fitted with appropriate peak functions where only the width was assumed to be constant for each population.⁴² Figure 5 shows the relative fractions of the three subpopulations for all six variants from the resulting peak areas (Figure 4). Overall, the intermediate state population is observed consistently between ~1 M and ~2.5 M GdmCl. The transition from the intermediate to the unfolded state occurs between ~2 and 2.5 M GdmCl for the different variants (Figure 5 and Figure S2) and coincides approximately with the unfolding midpoint concentration determined by CD spectroscopy (Figure 3b and d), confirming that the intermediate is still highly helical. At 1.75 M GdmCl, where the intermediate is populated almost exclusively according to the single-molecule experiments, it contains about 75% of the helicity of the native state (Figure 3b), in keeping with all helices except α A and α G being folded.

The data were fitted with an equilibrium three-state model (see Methods). Only for variant 107/157 (Figure 5b), the intermediate is not apparent, so a two-state model was used for the fit. In most cases, the data are reasonably well described by the equilibrium models. For variants 2/252 and 2/303, the unfolded state population exhibits a plateau between 1.25 and 2.25 M GdmCl, which is not well described by the fit (Figure 5e, f). A possible explanation is that the low transfer efficiency population in Figure 4e, f in this GdmCl concentration range may not correspond to the unfolded state but to a second intermediate conformation, which is in slow exchange with the first one, possibly linked to the unfolding/undocking of helix α A2. We note that the introduction of cysteine residues and labels at different positions of the polypeptide chain is likely to affect the stability of native state and intermediates to slightly different extents, which complicates a precise quantitative comparison of stabilities.

Chain Dynamics in the Intermediate State. Further evidence for the unfolding of helix αA in the intermediate comes from nanosecond FCS (nsFCS)⁶⁵ at 1.75 M GdmCl for variants 2/56 and 56/252, which probes the distance dynamics between donor and acceptor dyes (Figure 6). Besides the expected anticorrelated signal for very short lag times due to photon antibunching,⁶⁶ ClyA 2/56 shows an additional anticorrelated component in the donor-acceptor crosscorrelation and correlated components in the donor and acceptor autocorrelation curves, characteristic of fast distance dynamics in an unfolded chain (Figure 6a). A global fit of all three correlation curves yields a reconfiguration time of 40 ± 2 ns, in the same range as previously reported values for reconfiguration dynamics of unfolded chain segments of similar length ($\sim 10-100$ ns).^{67,68} In contrast, for variant 56/ 252, which probes the core of the protein, this additional component is not present in the nsFCS (Figure 6b), in agreement with the assumption that the core remains largely folded and hence rigid in the intermediate. The small correlated component in the donor autocorrelation is likely to originate from rotational dynamics, as supported by the asymmetry of the signal for positive and negative lag times.⁶⁹ To account for this rotational contribution, an additional component besides antibunching was included in the fit. (However, with an amplitude of 0.04, this component barely contributes to the fit of the donor-acceptor cross-correlation when compared to the amplitude of -0.34 for the anticorrelated signal characteristic of distance dynamics in the donor-acceptor cross-correlation of ClyA 2/56.) This observation is consistent with the 56/252 segment being largely folded and therefore rigid on the nanosecond time scale.

Finally, the dimensions of the intermediate were investigated by 2f-FCS⁴⁷ and compared to those of the native and unfolded states (Figure S3). The Stokes radius of the intermediate was found to be 3.5 ± 0.4 nm, not much greater than the value observed for the native state (3.2 ± 0.3 nm), in agreement with the presence of a folded helical core. Only upon complete unfolding does the Stokes radius increase substantially, reaching 5.9 ± 0.6 nm at 4 M GdmCl.

DISCUSSION AND CONCLUSION

Altogether, single-molecule FRET experiments unequivocally reveal the presence of an equilibrium unfolding intermediate for the cytolytic pore toxin ClyA. From the analysis of the unfolding experiments monitored by single-molecule FRET of six variants probing different protein segments, combined with CD, nsFCS, and 2f-FCS, the following picture emerges for the unfolding of ClyA (Figure 7): In the absence of denaturant,



Figure 7. Cartoon representation of the states involved in equilibrium folding of ClyA. In the absence of denaturant, only the native state (N) is populated. In the presence of $\sim 1-3$ M GdmCl, an intermediate state (I) is formed, in which the C- and N-terminal helices are unfolded. Increasing the denaturant concentration further leads to global unfolding of the protein (U).

only the native state (N) is populated. Addition of moderate concentrations of GdmCl (\sim 1–2.5 M) leads to the formation of an intermediate state (I) in which the C-terminal helix αG and a large part of the N-terminal helix αA are unfolded. This conclusion is based in particular on variant 2/56, which shows helix αA to be predominantly unfolded above 1.5 M GdmCl according to the change in FRET efficiency and the pronounced distance dynamics in nsFCS experiments. The loss of interactions between helices αA and αG is likely to contribute to the destabilization of the C-terminal helix αG_{i} which is unfolded in the intermediate state according to the properties of ClyA 187/290. This variant reveals an intermediate state population at very low transfer efficiency, indicating an expanded conformation of this segment. The three core helices αB , αC , and αF , as probed with variant 56/ 252, largely remain folded in the intermediate, since the native and intermediate state populations appear at similar transfer efficiencies. Additionally, this variant is devoid of nanosecond distance dynamics, as expected for a folded segment of the polypeptide chain. This interpretation is consistent with the results from CD spectroscopy, which show a helical content of \sim 75% in the range of denaturant concentrations where the intermediate is most populated.

The properties of the unfolding intermediate of ClyA may be linked to its biological function and the large conformational rearrangement of the soluble monomer into the protomer upon interaction with membranes or detergents. As part of this rearrangement, the N-terminal helix αA is translocated within the structure and rotated by almost 180° (Figure 1).³² This major conformational change requires a substantial amount of flexibility in the structure, presumably making helix αA prone to unfolding in the presence of relatively low concentrations of denaturant. Although helix αG is not affected as much during protomer conformation, it is shortened from both ends and loses its tight interactions with αA , likely to increase its tendency to unfold.

These local differences in stability are also reflected by a structural analysis of the differences between monomer and protomer conformation of ClyA. A map of side chain interactions allows us to distinguish contacts that are unique to monomer or protomer from those that are common to both conformations (Figure 1c). Unaltered interactions (displayed in black in Figure 1c) are primarily located at the interface between helices αB and αC . Interestingly, some contacts of helix αF with αG and αC are present in both conformations, even though the helical packing between helices αF and αC differs in the monomer and the protomer. Similarly, an alignment of the monomer and protomer structures illustrates the parts of the protein that are structurally most similar in both conformations (Figure 1c, inset). Helices αB and αC are hardly affected by the conformational transition, reflecting the stable core of ClyA. The largest change is apparent for helix αA_{i} as expected considering its large translocation during protomer formation.

The conformational bistability of ClyA may thus be encoded by differences in local stability across the protein that facilitate the rearrangement of parts of the structure, while others remain stably folded (Figure 1c). It would be interesting to use recent advances in coarse-grained models for the conformational changes of ClyA⁴¹ and for interpreting single-molecule FRET experiments⁷⁰ to develop a more detailed structural description of the intermediate. Such coarse-grained simulations have been previously applied to resolve the conformational switching of lymphotactin, another protein that can adopt two native states.⁷¹ It was shown that the transition involves a partly unfolded state that is stabilized by common contacts of both folds.⁷²

Could the equilibrium intermediate observed here be related to the kinetic intermediate observed previously in timeresolved single-molecule and CD measurements of ClyA?³⁰ In that work, an off-pathway intermediate was identified that is transiently populated during the interconversion from monomer to protomer upon interaction with the membranemimicking detergent DDM. However, the kinetic intermediate exhibited a very low transfer efficiency for variant 56/252 used in the time-resolve measurements, indicating that it is largely unfolded and can therefore not resemble the equilibrium intermediate observed here. It is conceivable, however, that an additional kinetic intermediate similar to the equilibrium intermediate identified here may exist, since its transfer efficiency for variant 56/252 is between those of the monomer and the protomer and may thus have eluded detection in the kinetics.⁴¹ Using the different variants employed here for the equilibrium measurements in the DDM-induced conversion kinetics may help to settle this question.

S Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpcb.8b07026.

Transfer efficiency histograms of different variants; dimensions of the native, intermediate, and unfolded state determined with dual-focus FCS (PDF)

AUTHOR INFORMATION

Corresponding Author

*Phone: +41 44 63 55535. E-mail: schuler@bioc.uzh.ch.

ORCID 6

Benjamin Schuler: 0000-0002-5970-4251

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank Rudi Glockshuber and Daniel Roderer for the pClyA plasmid and expression protocols and the Functional Genomics Center Zurich for expert mass spectrometry analysis. This work was supported by the Swiss National Science Foundation.

REFERENCES

(1) Eaton, W. A.; Muñoz, V.; Hagen, S. J.; Jas, G. S.; Lapidus, L. J.; Henry, E. R.; Hofrichter, J. Fast kinetics and mechanisms in protein folding. *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 327–59.

(2) Fersht, A. Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding; W. H. Freeman: 1999.

(3) Thirumalai, D.; O'Brien, E. P.; Morrison, G.; Hyeon, C. Theoretical perspectives on protein folding. *Annu. Rev. Biophys.* **2010**, 39, 159–83.

(4) Onuchic, J. N.; Wolynes, P. G. Theory of protein folding. *Curr. Opin. Struct. Biol.* **2004**, *14* (1), 70–75.

(5) Dill, K. A.; MacCallum, J. L. The protein-folding problem, 50 years on. *Science* **2012**, 338 (6110), 1042–6.

(6) Daggett, V. Protein folding-simulation. Chem. Rev. 2006, 106 (5), 1898–916.

(7) Jaenicke, R. Folding and association of proteins. Prog. Biophys. Mol. Biol. 1987, 49 (2-3), 117-237.

(8) Kim, P. S.; Baldwin, R. L. Specific intermediates in the folding reactions of small proteins and the mechanism of protein folding. *Annu. Rev. Biochem.* **1982**, *51*, 459–89.

(9) Schuler, B.; Eaton, W. A. Protein folding studied by singlemolecule FRET. *Curr. Opin. Struct. Biol.* **2008**, *18* (1), 16–26.

(10) Zoldak, G.; Rief, M. Force as a single molecule probe of multidimensional protein energy landscapes. *Curr. Opin. Struct. Biol.* **2013**, 23 (1), 48–57.

(11) Borgia, A.; Williams, P. M.; Clarke, J. Single-molecule studies of protein folding. *Annu. Rev. Biochem.* **2008**, *77*, 101–25.

(12) Banerjee, P. R.; Deniz, A. A. Shedding light on protein folding landscapes by single-molecule fluorescence. *Chem. Soc. Rev.* 2014, 43 (4), 1172–1188.

(13) Michalet, X.; Weiss, S.; Jäger, M. Single-molecule fluorescence studies of protein folding and conformational dynamics. *Chem. Rev.* **2006**, *106* (5), 1785–1813.

(14) Jefferson, R. E.; Min, D.; Corin, K.; Wang, J. Y.; Bowie, J. U. Applications of Single-Molecule Methods to Membrane Protein Folding Studies. *J. Mol. Biol.* **2018**, 430 (4), 424–437.

(15) Wang, Z.; Campos, L. A.; Munoz, V. Single-Molecule Fluorescence Studies of Fast Protein Folding. *Methods Enzymol.* **2016**, 581, 417–459.

(16) Gotz, M.; Wortmann, P.; Schmid, S.; Hugel, T. A Multicolor Single-Molecule FRET Approach to Study Protein Dynamics and Interactions Simultaneously. *Methods Enzymol.* **2016**, *581*, 487–516. (17) Fink, A. L. Compact Intermediate States in Protein Folding. *Annu. Rev. Biophys. Biomol. Struct.* **1995**, 24 (1), 495–522.

(18) Goto, Y.; Calciano, L. J.; Fink, A. L. Acid-induced folding of proteins. *Proc. Natl. Acad. Sci. U. S. A.* **1990**, 87 (2), 573–577.

(19) Ptitsyn, O. B.; Pain, R. H.; Semisotnov, G. V.; Zerovnik, E.; Razgulyaev, O. I. Evidence for a molten globule state as a general intermediate in protein folding. *FEBS Lett.* **1990**, *262* (1), 20–24.

(20) Korzhnev, D. M.; Religa, T. L.; Banachewicz, W.; Fersht, A. R.; Kay, L. E. A Transient and Low-Populated Protein-Folding Intermediate at Atomic Resolution. *Science* **2010**, *329* (5997), 1312–1316.

(21) Haran, G. Single-molecule fluorescence spectroscopy of biomolecular folding. *J. Phys.: Condens. Matter* **2003**, *15* (32), R1291–R1317.

(22) Banerjee, P. R.; Deniz, A. A. Shedding light on protein folding landscapes by single-molecule fluorescence. *Chem. Soc. Rev.* 2014, 43 (4), 1172–1188.

(23) Schuler, B.; Hofmann, H. Single-molecule spectroscopy of protein folding dynamics—expanding scope and timescales. *Curr. Opin. Struct. Biol.* **2013**, 23 (1), 36–47.

(24) Brucale, M.; Schuler, B.; Samori, B. Single-molecule studies of intrinsically disordered proteins. *Chem. Rev.* **2014**, *114* (6), 3281–317.

(25) 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 (7353), 662–5.

(26) Borgia, A.; Kemplen, K. R.; Borgia, M. B.; Soranno, A.; Shammas, S.; Wunderlich, B.; Nettels, D.; Best, R. B.; Clarke, J.; Schuler, B. Transient misfolding dominates multidomain protein folding. *Nat. Commun.* **2015**, *6*, 8861.

(27) Rhoades, E.; Gussakovsky, E.; Haran, G. Watching proteins fold one molecule at a time. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100* (6), 3197–202.

(28) Gambin, Y.; Schug, A.; Lemke, E. A.; Lavinder, J. J.; Ferreon, A. C. M.; Magliery, T. J.; Onuchic, J. N.; Deniz, A. A. Direct single-molecule observation of a protein living in two opposed native structures. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106* (25), 10153–10158.

(29) Pirchi, M.; Ziv, G.; Riven, I.; Cohen, S. S.; Zohar, N.; Barak, Y.; Haran, G. Single-molecule fluorescence spectroscopy maps the folding landscape of a large protein. *Nat. Commun.* **2011**, *2*, 493.

(30) Benke, S.; Roderer, D.; Wunderlich, B.; Nettels, D.; Glockshuber, R.; Schuler, B. The assembly dynamics of the cytolytic pore toxin ClyA. *Nat. Commun.* **2015**, *6*, 6198.

(31) Roderer, D.; Glockshuber, R. Assembly mechanism of the α -pore–forming toxin cytolysin A from *Escherichia coli*. *Philos. Trans. R. Soc., B* **2017**, 372 (1726), 20160211.

(32) Mueller, M.; Grauschopf, U.; Maier, T.; Glockshuber, R.; Ban, N. The structure of a cytolytic alpha-helical toxin pore reveals its assembly mechanism. *Nature* **2009**, *459* (7247), 726–30.

(33) Eifler, N.; Vetsch, M.; Gregorini, M.; Ringler, P.; Chami, M.; Philippsen, A.; Fritz, A.; Müller, S. A.; Glockshuber, R.; Engel, A.; Grauschopf, U. Cytotoxin ClyA from *Escherichia coli* assembles to a 13-meric pore independent of its redox-state. *EMBO J.* **2006**, 25 (11), 2652–2661.

(34) Soskine, M.; Biesemans, A.; De Maeyer, M.; Maglia, G. Tuning the Size and Properties of ClyA Nanopores Assisted by Directed Evolution. J. Am. Chem. Soc. **2013**, 135 (36), 13456–13463.

(35) Tzokov, S. B.; Wyborn, N. R.; Stillman, T. J.; Jamieson, S.; Czudnochowski, N.; Artymiuk, P. J.; Green, J.; Bullough, P. A. Structure of the Hemolysin E (HlyE, ClyA, and SheA) Channel in Its Membrane-bound Form. *J. Biol. Chem.* **2006**, *281* (32), 23042– 23049.

(36) Fahie, M. A.; Liang, L.; Avelino, A. R.; Pham, B.; Limpikirati, P.; Vachet, R. W.; Chen, M. Disruption of the open conductance in the β -tongue mutants of Cytolysin A. *Sci. Rep.* **2018**, *8* (1), 3796.

(37) Vaidyanathan, M. S.; Sathyanarayana, P.; Maiti, P. K.; Visweswariah, S. S.; Ayappa, K. G. Lysis dynamics and membrane

(38) Sathyanarayana, P.; Maurya, S.; Behera, A.; Ravichandran, M.; Visweswariah, S. S.; Ayappa, K. G.; Roy, R. Cholesterol promotes Cytolysin A activity by stabilizing the intermediates during pore formation. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115* (31), E7323.

(39) Wallace, A. J.; Stillman, T. J.; Atkins, A.; Jamieson, S. J.; Bullough, P. A.; Green, J.; Artymiuk, P. J. E. coli Hemolysin E (HlyE, ClyA, SheA): X-Ray Crystal Structure of the Toxin and Observation of Membrane Pores by Electron Microscopy. *Cell* **2000**, *100* (2), 265–276.

(40) Lella, M.; Mahalakshmi, R. Metamorphic Proteins: Emergence of Dual Protein Folds from One Primary Sequence. *Biochemistry* **2017**, 56 (24), 2971–2984.

(41) Giri Rao, V. V.; Desikan, R.; Ayappa, K. G.; Gosavi, S. Capturing the Membrane-Triggered Conformational Transition of an alpha-Helical Pore-Forming Toxin. *J. Phys. Chem. B* **2016**, *120* (47), 12064–12078.

(42) Dingfelder, F.; Wunderlich, B.; Benke, S.; Zosel, F.; Zijlstra, N.; Nettels, D.; Schuler, B. Rapid Microfluidic Double-Jump Mixing Device for Single-Molecule Spectroscopy. J. Am. Chem. Soc. **2017**, 139 (17), 6062–6065.

(43) Müller, B. K.; Zaychikov, E.; Bräuchle, C.; Lamb, D. C. Pulsed Interleaved Excitation. *Biophys. J.* **2005**, *89* (5), 3508–3522.

(44) Helenius, A.; McCaslin, D. R.; Fries, E.; Tanford, C. [63] Properties of detergents. In *Methods Enzymology*; Academic Press: 1979; Vol. 56, pp 734–749, DOI: 10.1016/0076-6879(79)56066-2.

(45) Schuler, B. Application of Single Molecule Förster Resonance Energy Transfer to Protein Folding. In *Protein Folding Protocols*; Bai, Y., Nussinov, R., Eds.; Humana Press: Totowa, NJ, 2006; pp 115– 138.

(46) Hofmann, H.; Soranno, A.; Borgia, A.; Gast, K.; Nettels, D.; Schuler, B. Polymer scaling laws of unfolded and intrinsically disordered proteins quantified with single-molecule spectroscopy. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109* (40), 16155–16160.

(47) Pieper, C.; Weiß, K.; Gregor, I.; Enderlein, J., Chapter Eight -Dual-Focus Fluorescence Correlation Spectroscopy. In *Methods Enzymology*; Tetin, S. Y., Ed.; Academic Press: 2013; Vol. 518, pp 175–204, DOI: 10.1016/B978-0-12-388422-0.00008-X.

(48) Nettels, D.; Hoffmann, A.; Schuler, B. Unfolded Protein and Peptide Dynamics Investigated with Single-Molecule FRET and Correlation Spectroscopy from Picoseconds to Seconds. *J. Phys. Chem. B* **2008**, *112* (19), 6137–6146.

(49) Gopich, I. V.; Nettels, D.; Schuler, B.; Szabo, A. Protein dynamics from single-molecule fluorescence intensity correlation functions. J. Chem. Phys. 2009, 131 (9), 095102.

(50) Zheng, W.; Zerze, G. H.; Borgia, A.; Mittal, J.; Schuler, B.; Best, R. B. Inferring properties of disordered chains from FRET transfer efficiencies. *J. Chem. Phys.* **2018**, *148* (12), 123329.

(51) Kalinin, S.; Peulen, T.; Sindbert, S.; Rothwell, P. J.; Berger, S.; Restle, T.; Goody, R. S.; Gohlke, H.; Seidel, C. A. M. A toolkit and benchmark study for FRET-restrained high-precision structural modeling. *Nat. Methods* **2012**, *9*, 1218.

(52) Gopich, I. V.; Szabo, A. Theory of the energy transfer efficiency and fluorescence lifetime distribution in single-molecule FRET. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109* (20), 7747–7752.

(53) Best, R. B.; Hofmann, H.; Nettels, D.; Schuler, B. Quantitative Interpretation of FRET Experiments via Molecular Simulation: Force Field and Validation. *Biophys. J.* **2015**, *108* (11), 2721–31.

(54) Petrášek, Z.; Schwille, P. Precise Measurement of Diffusion Coefficients using Scanning Fluorescence Correlation Spectroscopy. *Biophys. J.* **2008**, *94* (4), 1437–1448.

(55) Walters, J.; Milam, S. L.; Clark, A. C. Chapter 1 Practical Approaches to Protein Folding and Assembly: Spectroscopic Strategies in Thermodynamics and Kinetics. In *Methods Enzymology*; Academic Press: 2009; Vol. 455, pp 1–39, DOI: 10.1016/S0076-6879(08)04201-8.

(56) Dimura, M.; Peulen, T. O.; Hanke, C. A.; Prakash, A.; Gohlke, H.; Seidel, C. A. M. Quantitative FRET studies and integrative modeling unravel the structure and dynamics of biomolecular systems. *Curr. Opin. Struct. Biol.* **2016**, *40*, 163–185.

(57) Vehlow, C.; Stehr, H.; Winkelmann, M.; Duarte, J. M.; Petzold, L.; Dinse, J.; Lappe, M. CMView: Interactive contact map visualization and analysis. *Bioinformatics* **2011**, 27 (11), 1573–1574. (58) DeLano, W. L. *PyMOL molecular graphics system*; 2002. http://www.pymol.org (accessed Aug 31, 2018).

(59) Hansen, S.; Ernst, P.; König, S. L. B.; Reichen, C.; Ewald, C.; Nettels, D.; Mittl, P. R. E.; Schuler, B.; Pluckthun, A. Curvature of designed armadillo repeat proteins allows modular peptide binding. *J. Struct. Biol.* **2018**, *201* (2), 108–117.

(60) Andrade, M. A.; Chacón, P.; Merelo, J. J.; Morán, F. Evaluation of secondary structure of proteins from UV circular dichroism spectra using an unsupervised learning neural network. *Protein Eng., Des. Sel.* **1993**, *6* (4), 383–390.

(61) Whitmore, L.; Wallace, B. A. DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic Acids Res.* **2004**, 32 (suppl_2), W668– W673.

(62) Greenfield, N. J. Using circular dichroism spectra to estimate protein secondary structure. *Nat. Protocols* **2007**, *1*, 2876.

(63) Kouza, M.; Li, M. S.; O'Brien, E. P., Jr.; Hu, C. K.; Thirumalai, D. Effect of finite size on cooperativity and rates of protein folding. *J. Phys. Chem. A* **2006**, *110* (2), 671–676.

(64) Henry, E. R.; Hofrichter, J. [8] Singular value decomposition: Application to analysis of experimental data. In *Methods Enzymology*; Academic Press: 1992; Vol. 210, pp 129–192, DOI: 10.1016/0076-6879(92)10010-B.

(65) Schuler, B.; Hofmann, H.; Soranno, A.; Nettels, D. Singlemolecule FRET spectroscopy and the polymer physics of unfolded and intrinsically disordered proteins. *Annu. Rev. Biophys.* **2016**, *45*, 207–231.

(66) Plakhotnik, T.; Donley, E. A.; Wild, U. P. Single-molecule spectroscopy. *Annu. Rev. Phys. Chem.* **1997**, *48*, 181–212.

(67) Nettels, D.; Gopich, I. V.; Hoffmann, A.; Schuler, B. Ultrafast dynamics of protein collapse from single-molecule photon statistics. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104* (8), 2655–2660.

(68) Soranno, A.; Holla, A.; Dingfelder, F.; Nettels, D.; Makarov, D. E.; Schuler, B. Integrated view of internal friction in unfolded proteins from single-molecule FRET, contact quenching, theory, and simulations. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114* (10), E1833–E1839.

(69) Loman, A.; Gregor, I.; Stutz, C.; Mund, M.; Enderlein, J. Measuring rotational diffusion of macromolecules by fluorescence correlation spectroscopy. *Photochem. Photobiol. Sci.* **2010**, *9* (5), 627–36.

(70) Reinartz, I.; Sinner, C.; Nettels, D.; Stucki-Buchli, B.; Stockmar, F.; Panek, P. T.; Jacob, C. R.; Nienhaus, G. U.; Schuler, B.; Schug, A. Simulation of FRET dyes allows quantitative comparison against experimental data. *J. Chem. Phys.* **2018**, *148* (12), 123321.

(71) Tuinstra, R. L.; Peterson, F. C.; Kutlesa, S.; Elgin, E. S.; Kron, M. A.; Volkman, B. F. Interconversion between two unrelated protein folds in the lymphotactin native state. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105* (13), 5057–5062.

(72) Camilloni, C.; Sutto, L. Lymphotactin: How a protein can adopt two folds. J. Chem. Phys. 2009, 131 (24), 245105.

Supporting Information

Mapping an Equilibrium Folding Intermediate of the Cytolytic Pore Toxin ClyA with Single-Molecule FRET

Fabian Dingfelder¹, Stephan Benke¹, Daniel Nettels¹, & Benjamin Schuler^{1,2}

¹Department of Biochemistry, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland ²Department of Physics, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland



Figure S1. Transfer efficiency histograms of the different variants do not change during the time course of a measurement (30 min.), indicating that the samples are equilibrated. (a-f) Single-molecule FRET measurements of the different ClyA variants investigated in the presence of 1.5 M GdmCl. The data were split in overlapping 10 min intervals as indicated in the figure legend, and the resulting transfer efficiency histograms were overlaid.



Figure S2: Midpoint concentrations for the five variants of ClyA that clearly show the population of the intermediates state. Shown are the midpoint concentrations for the transition from the native to the intermediate state (blue) and from the intermediate to the unfolded state (red). The average midpoint concentrations of all five variants are depicted as dashed lines. Note that due to a large uncertainty in the m-values of the individual transitions, an accurate determination of free energy changes was not feasible.



Figure S3: The dimensions of the intermediate are close to those of the native state. (a) Transfer efficiency histograms of ClyA 56/252 for conditions under which the native state (top panel), intermediate state (middle panel) and unfolded state (bottom panel) are populated. The data correspond to those displayed in Figure 4d. (b) Dual-focus FCS measurements recorded under the same conditions as the data shown in (a), with the GdmCl concentration indicated in each panel. Fits used to quantify the translational diffusion time used to calculate the Stokes radius are shown as solid lines (see Methods). (c) Stokes radii as a function denaturant concentration, indicating that the intermediate state (red circle) is compact, with a Stokes radius close to that of the native state (green circle). In contrast, the unfolded state (blue circle) is much more expanded. The error bars were calculated assuming a ± 5 % uncertainty in the interfocus distance (see Methods).