Single-molecule Force Spectroscopy Reveals the Calcium Dependence of the Alternative Conformations in the Native State of a $\beta\gamma$-Crystallin Protein*[S]

Zackary N. Scholl[‡1,2], **Qing Li**[§1], **Weitao Yang**[¶3], **and Piotr E. Marszalek**[§4]

From the ‡Program in Computational Biology and Bioinformatics, the §Department of Mechanical Engineering and Materials Science, and the ¶Department of Chemistry, Duke University, Durham, North Carolina 27708

Although multidomain proteins predominate the proteome of all organisms and are expected to display complex folding behaviors and significantly greater structural dynamics as compared with single-domain proteins, their conformational heterogeneity and its impact on their interaction with ligands are poorly understood due to a lack of experimental techniques. The multidomain calcium-binding $\beta\gamma$-crystallin proteins are particularly important because their deterioration and misfolding/aggregation are associated with melanoma tumors and cataracts. Here we investigate the mechanical stability and conformational dynamics of a model calcium-binding $\beta\gamma$-crystallin protein, Protein S, and elaborate on its interactions with calcium. We ask whether domain interactions and calcium binding affect Protein S folding and potential structural heterogeneity. Our results from single-molecule force spectroscopy show that the N-terminal (but not the C-terminal) domain is in equilibrium with an alternative conformation in the absence of Ca^{2+}, which is mechanically stable in contrast to other proteins that were observed to sample a molten globule under similar conditions. Mutagenesis experiments and computer simulations reveal that the alternative conformation of the N-terminal domain is caused by structural instability produced by the high charge density of a calcium binding site. We find that this alternative conformation in the N-terminal domain is diminished in the presence of calcium and can also be partially eliminated with a hitherto unrecognized compensatory mechanism that uses the interaction of the C-terminal domain to neutralize the electronegative site. We find that up to 1% of all identified multidomain calcium-binding proteins contain a similarly highly charged site and therefore may exploit a similar compensatory mechanism to prevent structural instability in the absence of ligand.

Multidomain proteins are highly prevalent in the proteomes of all organisms (1); however, there is still little known about their folding pathways and interaction with ligands (2), as compared with small, single-domain proteins. This is because large, multidomain proteins may exhibit significant structural dynamics and transient interactions between the domains that are rather difficult to capture and analyze. An important example of this class of proteins are the calcium-binding proteins that are effectors for stimulating cellular signals (3, 4) and buffering calcium levels (5). Currently, of particular interest are the calcium-binding $\beta\gamma$-crystallin proteins that are implicated in several human diseases (6–8). The complexity of these proteins may best be understood through single-molecule methods that are able to deconvolve their conformational heterogeneity. Although there have been exciting recent single-molecule studies of the multidomain calmodulin, including the impact of domains' interactions on calcium binding (9), and studies on mechanical unfolding of $\beta\gamma$-crystallins that captured their interesting domain swapping behavior (10), there has not yet been any single-molecule study of the calcium dependence on folding and mechanics for multidomain $\beta\gamma$-crystallins.

Here, we ask what is the mechanical stability of the domains of Protein S in isolation and how it is altered by the domain-domain interaction in wild type Protein S and upon binding calcium. To investigate, we use single-molecule force spectroscopy (SMFS)[5] using an atomic force microscope (reviewed in Refs. 11–16). SMFS has an unprecedented ability to distinguish between populations from a mixture of conformations and allows precise monitoring of changes in stability or structure during unfolding and refolding. SMFS and computer simulation is the method of choice for studying multidomain proteins because it minimizes protein aggregation and allows the study of the intrinsically asynchronous processes involved in folding (9, 17–20).

We chose to study the model protein, Protein S (Fig. 1), because it is the founding member of the $\beta\gamma$-crystallin superfamily (21, 22) and exhibits calcium-dependent folding (23). The $\beta\gamma$-crystallin superfamily is defined by proteins that have two domains, with each domain containing an eight-stranded β-sandwich composed of two similar "Greek key" motifs (22, 24, 25). The $\beta\gamma$-crystallins are found in all kingdoms of life and are commonly found at high protein concentrations and typically function to confer stress resistance (26). In humans, the $\beta\gamma$-crystallin proteins include AIM1, a protein associated with

[S] This article contains supplemental Table 1.
[1] Both authors contributed equally to this work.
[2] To whom correspondence may be addressed. E-mail: zns@duke.edu.
[3] To whom correspondence may be addressed. E-mail: weitao.yang@duke.edu.
[4] To whom correspondence may be addressed. E-mail: pemar@duke.edu.

[5] The abbreviations used are: SMFS, single-molecule force spectroscopy; PS, Protein S; NTD-PS and CTD-PS, N- and C-terminal domain truncations of PS, respectively; pN, piconewtons; AFM, atomic force microscopy.
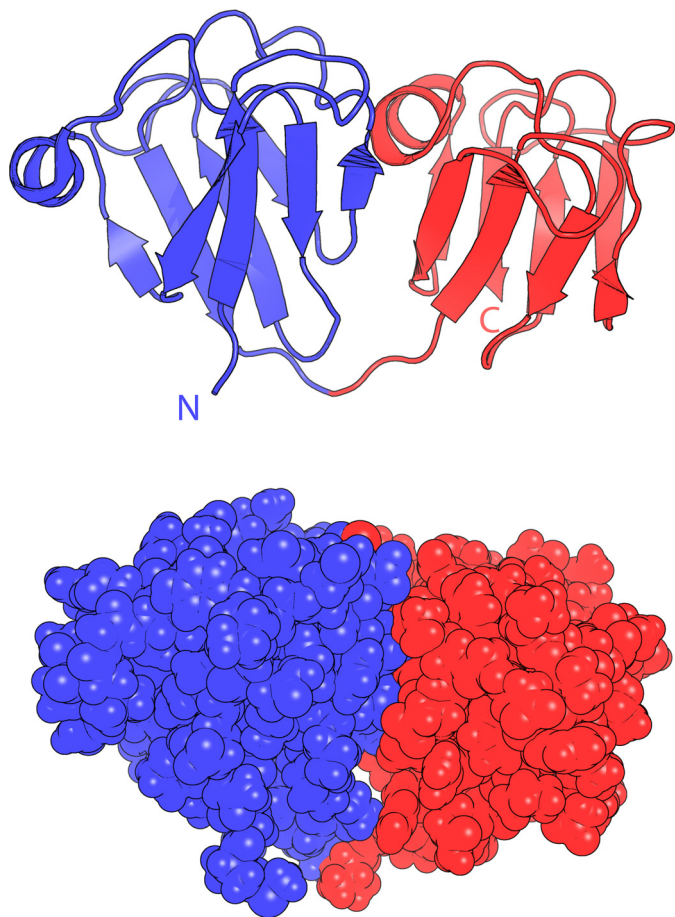
FIGURE 1. **Schematic structure of Protein S** *colored* **from N terminus (*blue*) to C terminus (*red*) in schematic representation (*top*) and all-atom sphere representation (*bottom*).**

melanoma tumors (6), as well as many different eye lens proteins (27) whose misfolding has been implicated in cortical cataracts (7, 8). It is well known that the $\beta\gamma$-crystallins bind calcium (28–30), and without calcium, the $\beta\gamma$-crystallins are often found to be less stable (31–34) or unstructured (35). However, there is little known about conformational heterogeneity of the $\beta\gamma$-crystallin proteins (36) and the role of calcium and multidomain interactions supporting their stability.

The previous ensemble studies on truncated domains of Protein S (23, 37–39) suggested that the domains are less stable in the absence of calcium, but the mechanistic origin of this instability has remained poorly understood. We confirmed these earlier observations using SMFS on isolated domains and then performed a novel single-molecule study of the complete multidomain Protein S that identifies the origin of the domains' instability and the resulting conformational heterogeneity in the absence of calcium as due to electrical properties of the structure surrounding the calcium binding site. We also identified a new "compensatory" mechanism that, in the absence of the ligand, helps to stabilize the weakened domains through specific interdomain interactions.

Results

Truncated N-terminal Domain of Protein S Is in Equilibrium with an Alternative Conformation in the Absence of Ca^{2+}—To understand how the Protein S domains interact in the multido-

main protein, we first studied the truncated N-terminal and truncated C-terminal domains. The N-terminal domain truncation (NTD-PS) was made, at the DNA level, by cloning the Protein S N-terminal domain containing residues 1–88 of Protein S into a construct such that it would be flanked on either side by three I27 domains from titin (40, 41). The I27 domains serve as a positive control of full unfolding of the NTD-PS molecule (as long as at least four I27 domains are unfolded in the force-extension profile) and also allow unequivocal detection of single-molecule events (if clear unfolding of I27 is seen, it is probably a single-molecule event and not a multiple-molecule recording).

The NTD-PS molecule was stretched along the N-C pulling geometry through tethering from the surface to the cantilever tip (Fig. 2). A representative force-extension profile of the unfolding of NTD-PS is shown in Fig. 2*A* in *red*, and the subsequent unfolding of the I27 domains is in *blue*. It can be seen that NTD-PS unfolding creates a single peak in the force-extension spectrum. The unfolding of I27 can be discerned from the unfolding of Protein S because the unfolding forces and contour length increments do not overlap (Fig. 2*B*). The crystal structure measurements indicate that the initial length of each domain is 1.0 nm. Given that an extended polypeptide provides 0.365 nm of extension (42), the unfolding of the N-terminal domain should theoretically provide a contour length increment of 0.365 nm/residue × 89 residues − 1.0 nm = 31.5 nm, which agrees with the measured contour length increment of 31.9 ± 3.4 nm (mean ± S.D.), also shown in Fig. 3*B*.

The unfolding forces of NTD-PS indicate that it is less stable than I27 (and hence unfolds first) because the peak typically unfolds between 40 and 120 pN, whereas I27 unfolds at 207 ± 28 pN (at 4,500 pN/s). Interestingly, the null hypothesis that the distribution of unfolding forces for the NTD-PS comes from a normal distribution is rejected ($p = 0.026$, Kolmogorov-Smirnov test for normality). The histogram of rupture forces ($n = 236$ total recordings) of NTD-PS is bimodal and can be fit by two normal distributions (Fig. 3*A*) with mean ± S.D. of 49 ± 14 pN (55 ± 6% of recordings) and 93 ± 14 pN (45 ± 6% of recordings).

Why is the force distribution of NTD-PS bimodal? It is possible that the two states could come from transitions induced by the pulling experiment itself (43). To test this hypothesis, we conducted experiments at a speed that was 15-fold faster than previous experiments under the assumption that the faster speed would decrease any likelihood of conformational changes during the pulling measurements. These experiments show that the bimodality still exists and that the proportions of states are not statistically significant from the proportions at the slower speed (χ^2 test, $p > 0.05$; Fig. 4), indicating that the pulling experiment itself has no effect on the proportions of the recordings with low or high force peaks. Although it is possible that we may need to reach higher pulling extremes to detect the pulling-induced transition, our interpretation is consistent with previous NMR experiments that identified slow exchanging intermediates in NTD-PS in equilibrium (39).

Because the bimodality is not an artifact of the pulling experiment, we hypothesized that the lower unfolding force distribu-
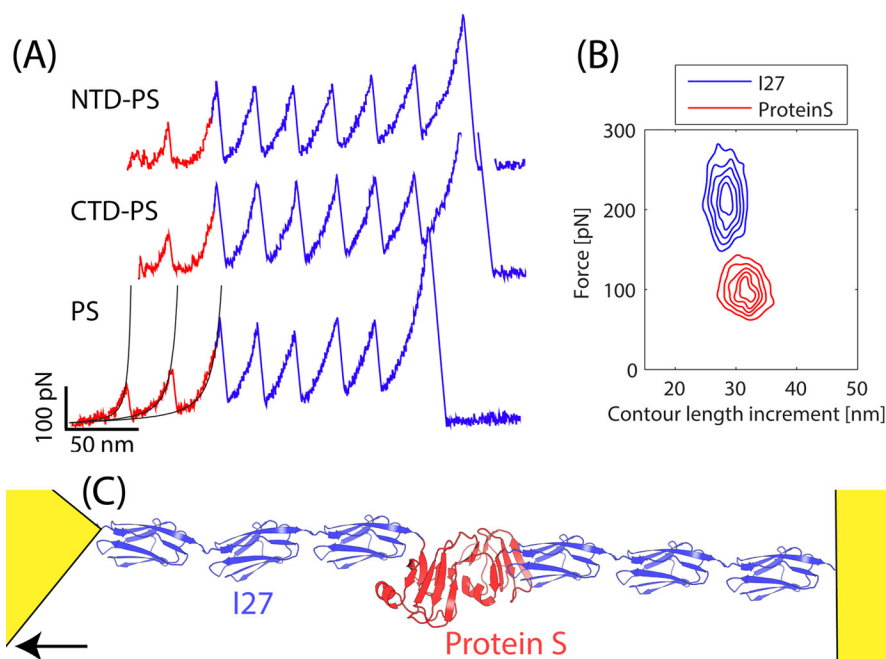
FIGURE 2. *A*, representative force-extension curves from the unfolding of I27-flanked NTD-PS, CTD-PS, and PS. *B*, probability density of peaks having a given force and contour length increment for I27 (*blue*) and Protein S (*red*). Protein S peaks can be readily discerned from I27 peaks due to their non-overlapping densities. *C*, schematic of the unfolding experiment of Protein S flanked by three I27 domains on either side.
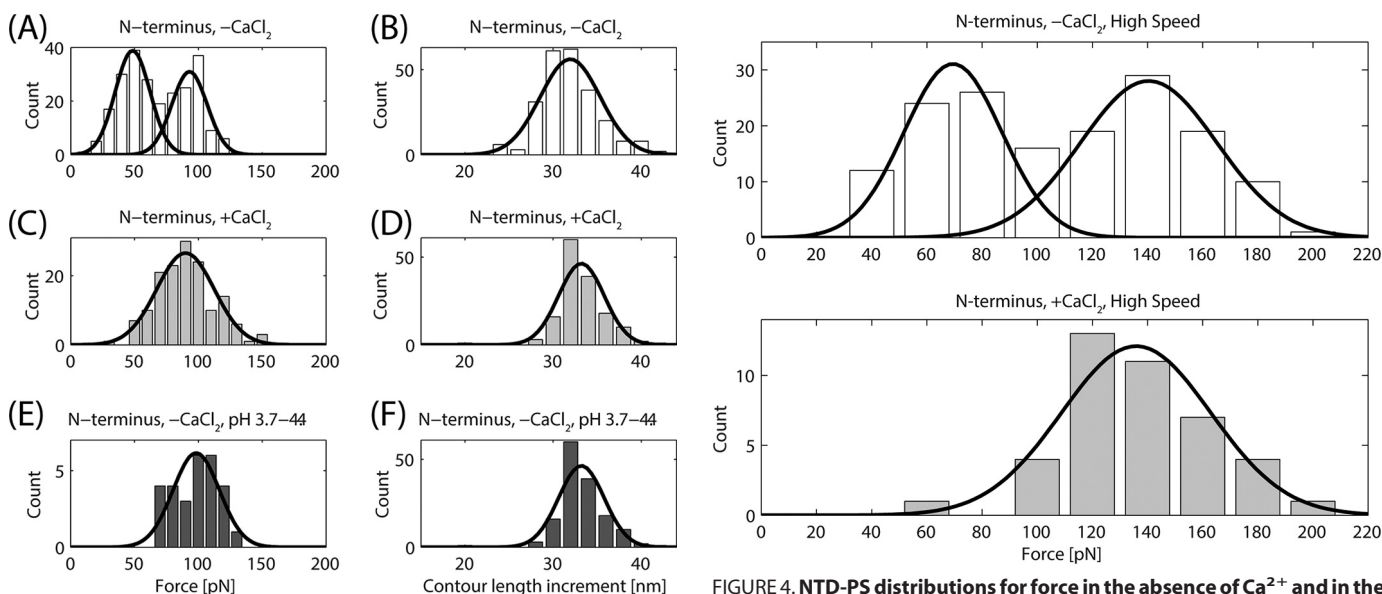

FIGURE 3. **NTD-PS distributions for force and contour length increment in the absence of Ca^{2+} (*A* and *B*), in the presence of Ca^{2+} (*C* and *D*), and in the absence of Ca^{2+} at low pH (*E* and *F*).**

FIGURE 4. **NTD-PS distributions for force in the absence of Ca^{2+} and in the presence of Ca^{2+} for pulling experiments done at a high loading rate (216,000 pN/s).** Approximately $46 \pm 8\%$ of recordings exhibiting a mechanically weaker unfolding event and $54 \pm 8\%$ exhibiting a mechanically stronger unfolding event. These proportions are slightly shifted from the experiments with lower loading rates but not statistically significant (χ^2 test, $p > 0.05$).

tion comes from the unfolding of an alternative conformation of NTD-PS that contains only a fraction of the contacts. This is supported by the fact that whereas force-distribution is multimodal, the contour length increment distribution is not (Fig. 3*B*). When the measured contour length increments are filtered based on the force, the mean $\pm$ S.D. of the contour length increment of the low force distribution (<71 pN) is 31.6 ± 2.7 nm, and the high force distribution (≥ 71 pN) is 31.8 ± 2.6 nm, which are not statistically significantly different ($p = 0.4$) and also not statistically significantly different from what would be expected for the full NTD-PS domain ($p = 0.6$). Because the

contour length increment is unchanged, the structural position of the N-terminal and C-terminal residues of the NTD-PS domain must be fixed. The fact that the mechanical stability is decreased in the alternative conformation would imply that contacts are missing in that state compared with the higher force state because contact density highly correlates with unfolding force (44, 45).

The $\beta\gamma$-crystallin proteins are often stabilized by Ca^{2+} (28–30), and it is known that Protein S can bind calcium (46),

although the location of the binding sites has not reached a consensus (37, 46). We hypothesized that the alternative conformation might be due to charge repulsion, which then could be converted to a single native configuration by the addition of Ca^{2+} or lowering pH. Interestingly, the addition of Ca^{2+} produced an unfolding distribution for NTD-PS that was not bimodal, which had a mean ± S.D. of 90 ± 23 pN (Fig. 3C). The effect of Ca^{2+} on the unfolding modality was also speed-independent (Fig. 4). Similarly, the presence of an acidic environment (pH 3.7–4.4) also has unfolding forces of NTD-PS in a single distribution at 99 ± 23 pN (Fig. 3D). In both of these cases, the contour length increment matches the unfolding of NTD-PS without $CaCl_2$ (Fig. 3, *B*, *D*, and *F*). Thus, the presence of Ca^{2+} or an acidic environment acts to convert the alternative conformation to a single conformation. These observations are also consistent with independent NMR experiments on the NTD-PS, which show that an intermediate is undetectable in conditions with $CaCl_2$ or an acidic pH (39).

Truncated C-terminal Domain of Protein S Has Higher Mechanical Stability as Compared with the Truncated N-terminal Domain—The C-terminal domain of Protein S is highly similar to the N-terminal domain with 55% sequence identity and a root mean square deviation between matched Cα atoms of 2.0 Å. Because of the high similarity, we hypothesized that the truncated C-terminal domain would also show bimodal distributions of force rupture events due to the presence of an also similar alternative conformation. An I27-flanked truncated C-terminal domain (CTD-PS) was used for pulling, and a representative force-extension curve of CTD-PS is shown in Fig. 2A.

Contrary to our hypothesis, the unfolding of CTD-PS in the absence of Ca^{2+} shows a force distribution that is unimodal and not bimodal (Fig. 5). The CTD-PS force distribution can be fit with a single normal distribution with mean and S.D. of 96 ± 21 pN. This stability is very similar to the high force conformation of the N-terminal domain of Protein S. The contour length increment is 31.7 ± 2.4, which is similar to the calculated contour length increment of the C-terminal domain of Protein S, indicating that the full protein unfolds. The C-terminal domain of Protein S also binds Ca^{2+}, which could increase stabilization. In the presence of Ca^{2+}, the unfolding force of CTD-PS increased to 115 ± 17 pN. Although the C-terminal domain has very high similarity to the N-terminal domain, it does not show evidence of a stable alternative conformation at equilibrium, with or without Ca^{2+}. Instead, Ca^{2+} directly stabilizes the native fold of the C-terminal domain. Interestingly, In some proteins, proline isomerization may introduce alternative conformations (47). Indeed, Protein S has 12 prolines that could assume isomerization states. However, because both domains have the same number of prolines but only NTD-PS adopts the alternative conformations, we hypothesize that this effect is not the origin of the alternative conformation. This will be further tested in future studies.

Protein S N-terminal Domain Adopts an Alternative Conformation in the Absence of Ca^{2+}—To study the extent to which the truncated domains of Protein S contribute to the mechanically stability of the full Protein S, we performed similar single-molecule force spectroscopy experiments on the full Protein S (PS) in the presence and absence of Ca^{2+}. Protein S was flanked by I27 domains and unfolded along the N-C coordinate by pulling at constant velocity. As can be seen from the representative force-extension curve shown in Fig. 2A, the unfolding occurs in two steps, which are denoted Peak 1 and Peak 2. The force distribution and contour length distributions for both peaks in the absence of Ca^{2+} are shown in Fig. 6 (*A–D*).
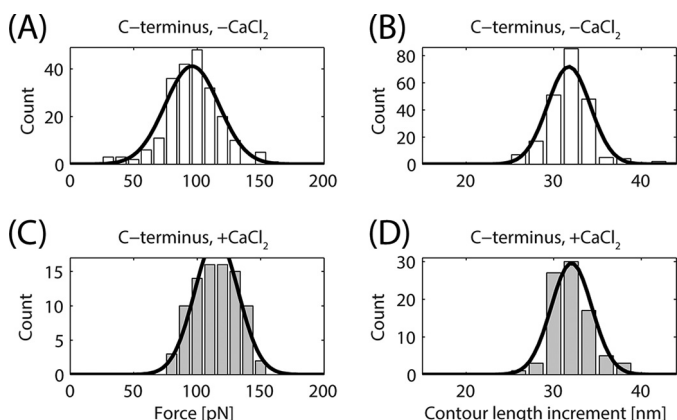

FIGURE 5. **CTD-PS distributions for force and contour length increment in the absence of Ca^{2+} (*A and B*) and in the presence of Ca^{2+} (*C and D*).**
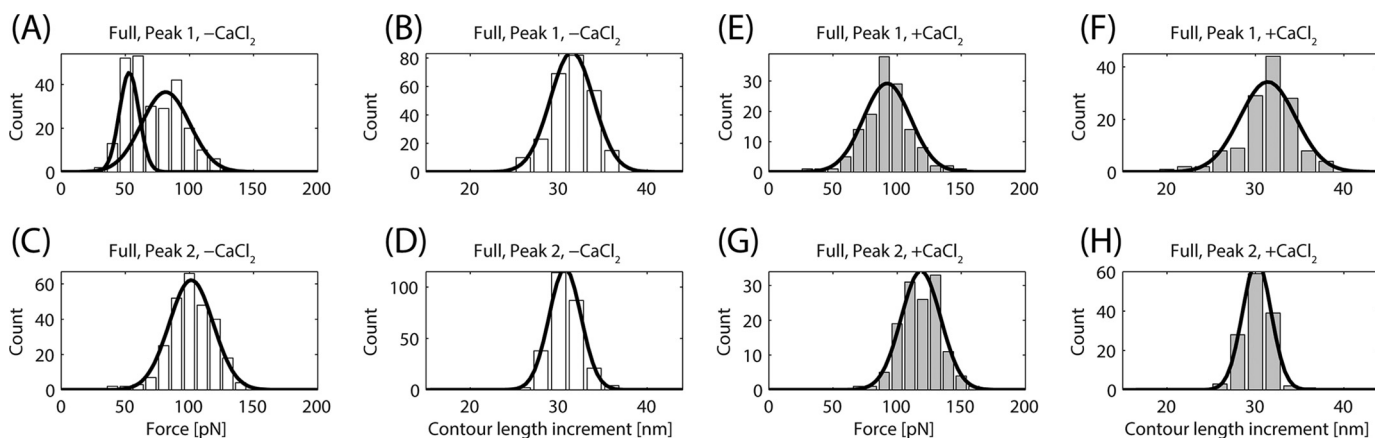

FIGURE 6. **PS distributions of force and contour length increment for both Peak 1 and Peak 2 in the absence of Ca^{2+} (*A–D*) and in the presence of Ca^{2+} (*E–H*).**
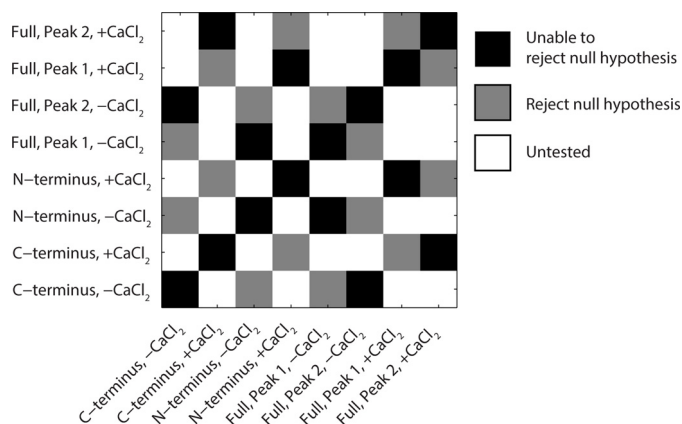
FIGURE 7. **Each pair of force distributions are paired and tested using the null hypothesis that the distributions come from the same continuous distribution (Kolmogorov-Smirnov test).** Pairs that reject the null hypothesis are shown in *gray*, and pairs that cannot reject the null hypothesis are shown in *black*. Pairs that do not have the $CaCl_2$ conditions are untested, shown in *white*. Similar results were obtained using the more stringent Anderson-Darling test for comparison.

FIGURE 8. **In the absence of calcium, there is a cluster of four negatively charged residues in the N-terminal domain of Protein S (Glu[10], Asp[11], Glu[70], and Glu[71]).** The C-terminal domain provides a lysine (Lys[104]) that may interact with this cluster in order to provide further stabilization in the absence of calcium.

(A) N-terminal domain

(B) C-terminal domain

FIGURE 9. **Charge distributions of the N-terminal domain (*A*) and C-terminal domain (*B*) shown in all-atom representation (*left*) and *zoomed in* on the residues in the structure (*right*).** All structural representations are aligned in the same orientation.

The contour length increments of Peak 1 and Peak 2 agree with the theoretical extension of the N-terminal or C-terminal domain; however, since the numbers of residues in each domain are nearly the same, the contour length increment cannot be used to accurately assign the unfolding of a particular domain to a peak in the force-extension curve. To assign Peak 1 and Peak 2 to the unfolding of a particular domain, we tested whether their mechanical stabilities matched that of the truncated domains. We tested each pair of distributions against one another (under the same condition of $CaCl_2$) using the null hypothesis that the distributions come from the same continuous distribution (Fig. 7). The null hypothesis was rejected if the p value was below a Bonferroni-corrected p value of 0.05. This analysis shows that the only cases in which the null hypothesis cannot be rejected are NTD-PS ($+CaCl_2$) and PS-1 ($+CaCl_2$), NTD-PS ($-CaCl_2$) and PS-1 ($-CaCl_2$), CTD-PS ($+CaCl_2$) and PS-2 ($+CaCl_2$), and finally CTD-PS ($-CaCl_2$) and PS-2 ($-CaCl_2$). This indicates that the stability of the truncated CTD-PS domain corresponds to the mechanical stability of the domain that gives rise to Peak 2 in Protein S, and the truncated NTD-PS domain has mechanical stability similar to that of the domain that gives rise to Peak 1 in Protein S. Thus, the mechanical unfolding of Protein S begins with the unfolding of the N-terminal domain and is followed by the unfolding of the C-terminal domain.

The mechanical stability of the N-terminal domain of Protein S (Peak 1), like the NTD-PS, has a multimodal force distribution for Peak 1 ($p = 0.009$, Kolmogorov-Smirnov test for normality). The force distribution for Peak 1 can be fit by two normal distributions with mean ± S.D. of 53 ± 8 pN and 82 ± 19 pN (Fig. 6*A*). The recordings with lower force were only a minor fraction of the total recordings, accounting for 34% of all of the recordings ($n = 258$). Also, similar to the NTD-PS, the presence of a minor fraction in the force distribution is absent when the experiments are done in a solution containing Ca^{2+} ions (Fig. 6*E*). These results indicate that the full Protein S, in the absence of Ca^{2+}, also contains an alternative conformation in the N-terminal domain, that has mechanical stability similar
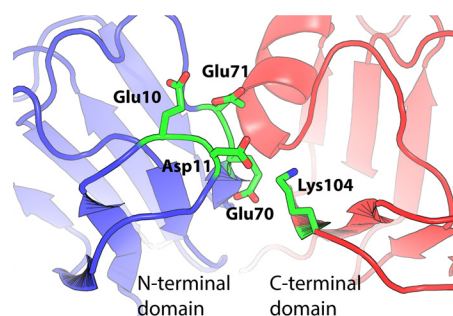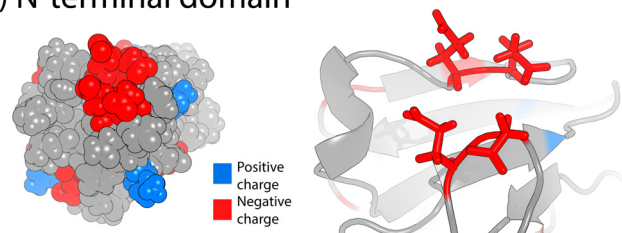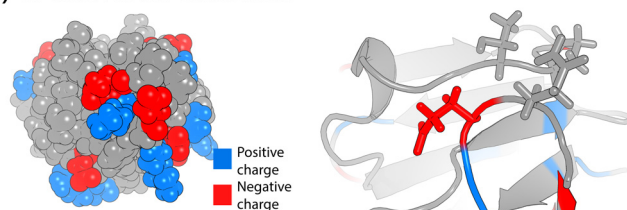
to that of the truncated N-terminal domain (Fig. 3*A*). Whereas the NTD-PS had 55% of recordings in the alternative conformation, the N-terminal domain in the full PS has only 34% of recordings in the alternative conformation, probably due to stabilization provided from the C-terminal domain. Indeed, the crystal structure shows that the C-terminal domain provides a lysine that interacts closely with the N-terminal domain (Fig. 8), which may act as a compensating mechanism for increasing stability in the absence of calcium.

Charge-destabilized Conformation of N-terminal Domain Confirmed by Molecular Dynamics Simulations and Mutagenesis—Why do experiments show that the truncated N-terminal domain of Protein S has a stable alternative conformation, whereas the highly homologous C-terminal domain does not? Although the structures have high similarity, one major difference between the domains is that the N-terminal domain has a cluster of four negatively charged residues, which is absent in the C-terminal domain (Fig. 9, *A* and *B*). Thus, we hypothesize that, in the absence of the C-terminal residues or in the absence of Ca^{2+} in the solution, these residues may repel and destabilize the loop, resulting in an alternative conformation where the N-terminal domain loops have a different configuration. The destabilization of these loops could decrease the

mechanical stability of the protein but would not affect the contour length increment because these residues are not part of the β strands that clinch the N-terminal and C-terminal residues of NTD-PS.

We used molecular dynamic simulations to investigate whether the cluster of charge residues (Fig. 9) in the N-terminal domain causes a destabilization of the loops in the truncated domain, without Ca^{2+}. We monitored the Glu^{10}–Glu^{70} distance in simulations of the truncated N-terminal domain of Protein S with and without Ca^{2+} and the N-terminal domain within the full Protein S without Ca^{2+}. To compare with the C-terminal domain, we also conducted a simulation of the truncated C-terminal domain and monitored the analogous loop distances between Gln^{100} and Pro^{158}. Simulations revealed that only the truncated N-terminal domain in the absence of the C-terminal domain and in the absence of Ca^{2+} showed a significant deviation in the loop stability (Fig. 10A).

In the simulation of the truncated N-terminal domain without Ca^{2+}, the Glu^{10} loop had been completely repelled away from the Glu^{70} loop (Fig. 10B). However, for all other simulations (truncated N-terminal domain with Ca^{2+}, truncated

C-terminal domain, or N-terminal domain in the full Protein S), there was no major conformational change between the loops. In each case, the loop was stabilized by an increased net charge. For instance, in the case of the truncated N-terminal domain with Ca^{2+}, we observed Ca^{2+} transiently binding to the charged residues, which prevented loop repulsion. Also, in the case of the full Protein S, we observed a C-terminal domain lysine (Lys^{104}) interacting with the charged residues of the N-terminal domain, again supporting the hypothesis that the multiple domains work in unison as a compensatory mechanism for promoting stability.

To experimentally verify the observation that the charged loop region is responsible for the alternative conformation, we generated a mutant Protein S with the negatively charged residues in the N-terminal domain converted to their C-terminal domain counterparts, namely D11Q, E70P, and E71T (Fig. 11). The mutations for disrupting the charge repulsion site where determined through alignment between the N-terminal domain of Protein S and the C-terminal domain of Protein S shown in Fig. 12. The highly charged site of the N-terminal domain (in *blue*) was then converted to the corresponding residues in the C-terminal domain (in *red*) so that the final Protein S mutant would have the D11Q, E70P, and E71T mutations. The sequence homology after the mutation increased from 55% to 59%.

We performed force spectroscopy on this mutant and measured the distribution of unfolding forces (Fig. 12). We observed that the mean unfolding force for Peak 2 was unaffected by this mutation (Student's t test, $p = 0.09$ and $p = 0.30$ with and without added Ca^{2+}, respectively), which indicates that Peak 2 probably corresponds to the unfolding of the C-terminal domain, as in wild type, and that the mutations on the N-terminal domain have no effect on the C-terminal domain. Interestingly, the unfolding force distribution for Peak 1 lost all bimodality without Ca^{2+}. This indicates that the residues 10–11 and 70–71 in the wild type Protein S are probably the source of the instability that leads to an alternative conformation. In future studies, we plan to also graft residues in the opposite way, to the C-terminal domain from the N-terminal domain, as well as testing mutated truncated variants.

Speed of Protein S Refolding Is Calcium-dependent—We examined the calcium dependence for refolding of the full Protein S using force spectroscopy cyclic pulse experiments (40, 48). Using a smaller extension, the molecule was stretched enough to unfold only Protein S, and upon retraction, the tip was moved 10–20 nm away from the surface to prevent surface interaction. Then subsequent extension/retraction cycles were enacted with a specified time delay before each pulse, which varied between 100 ms and 10 s (plus an additional 260 ms from
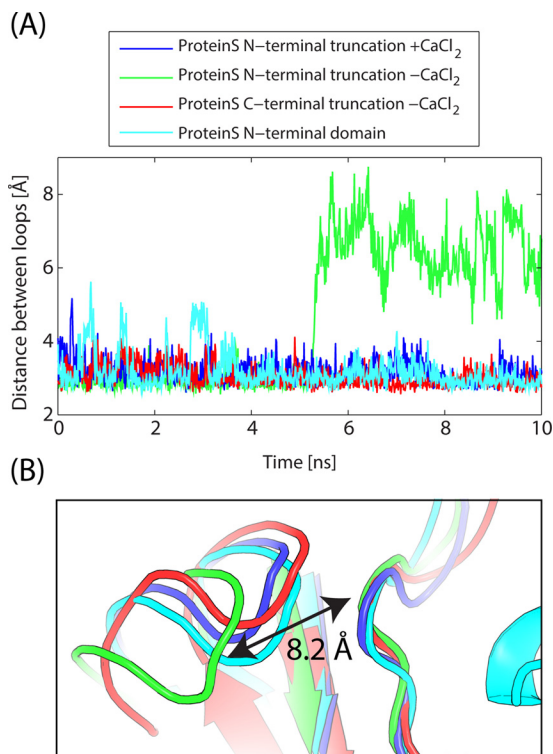

FIGURE 10. *A*, molecular dynamic time trace of the distance between two loops in Protein S (distance between Glu^{10} and Glu^{70} in the N-terminal domain and distance between Gln^{100} and Pro^{158} in the C-terminal domain). Without a C-terminal domain and without Ca^{2+}, the N-terminal domain (*green*) clearly has the two loops repelling from one another. *B*, snapshot of the loop regions of the four constructs (same *color code* as in *A*).

FIGURE 11. **Protein S domains sequence alignment.** For mutagenesis experiments, the *blue* residues of the N-terminal domain were mutated to the *red* residues of the C-terminal domain.
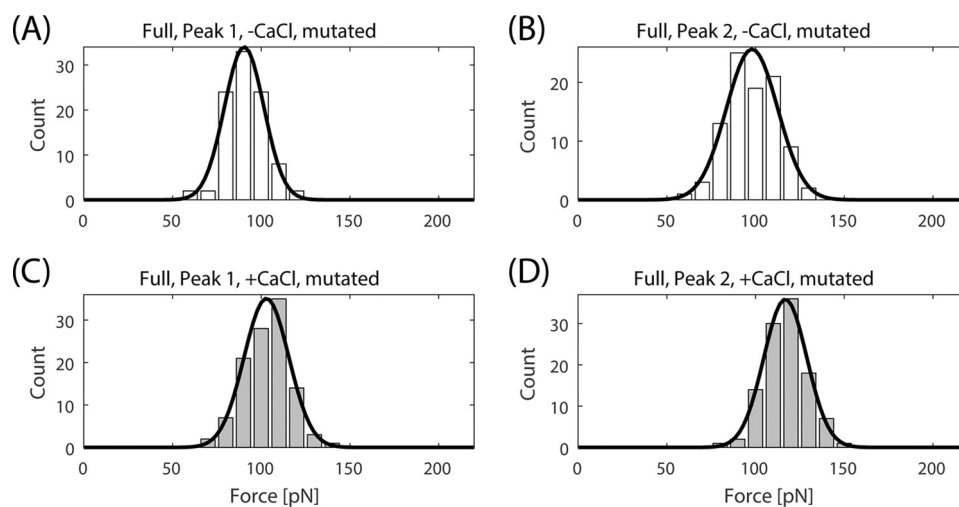
FIGURE 12. *A* and *B*, Protein S D11Q/E70P/E71T was pulled and the rupture force was measured, without calcium, for Peak 1 (*A*) and Peak 2 (*B*). *C* and *D*, the same measurements were done with calcium for Peak 1 (*C*) and Peak 2 (*D*).
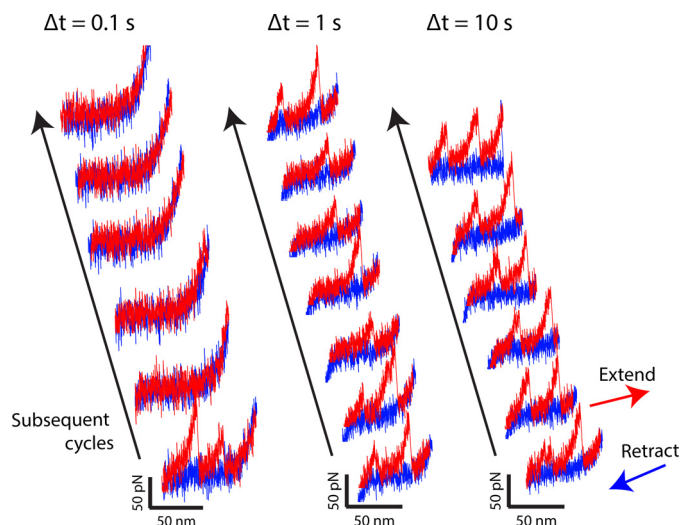

FIGURE 13. **Force-extension curves obtained from refolding experiments on Protein S.** The force-extension curve in *red* shows the extension, and the curve in *blue* shows the retraction for a particular molecule. Each *column* is a single experiment starting with the first extension and retraction (*bottom*) and then each subsequent extension and retraction after a designated amount of time (0.1, 1, or 10 s). The number of domains that appear increases with the time delay.

the dead time during relaxation). During this time delay, the protein can spontaneously refold. Refolding is detected by a subsequent unfolding pulse that will probe the conformation reached in the time interval. As can be seen in Fig. 13, the longest refolding time (10.3 s) is most likely to show refolding as judged by the force-extension curve resembling that of a natively folded protein. The smallest time delay corresponds to almost no fully folded proteins (no force peaks in the force-extension curve).

The presence of an alternative configuration from the refolding was quantified using the previously determined force distributions (Fig. 6). Any force rupture event that was at least 2 S.D. values from the mean of the previously determined force distributions of Peak 1 and Peak 2 was selected as an event in which the domain reached its native state. For Peak 1 without $CaCl_2$, the normal distribution with higher mean was used to determine the cutoff. The types of recovered events can then fall into

one of the following six categories defined by the number of peaks and the rupture force of each peak: 1) no unfolding events; 2) one low force event (defined by the previously characterized force distributions in Fig. 6); 3) one normal event; 4) two low force events; 5) one low force + one normal event; 6) two normal events. In this taxonomy, type 1 indicates "fully unfolded," whereas type 6 indicates "fully folded" and types 2–5 indicate alternative conformations that we refer to as "partially folded."

Refolding experiments were performed in the presence of or absence of $CaCl_2$ at five different total refolding times: 360 ms, 0.8 s, 1.3 s, 3.3 s, and 10.3 s. A pie chart for the taxonomy of recovered events is shown in Fig. 14. All six of the possible events were detected as shown in Fig. 14 with overlapped examples of the extension recordings after refolding (*red*) shown over examples of a normal Protein S unfolding without refolding (*blue*). These results show that the presence of $CaCl_2$ increases the rate of refolding because the number of fully folded Protein S molecules is much higher at shorter delays with the addition of $CaCl_2$.

Interestingly, there is a small fraction of events in conditions with and without Ca^{2+} that show two small peaks (number 4 in taxonomy, *cyan* in Fig. 14). This force-extension recording is only possible if both the N-terminal domain and the C-terminal domain adopt a less stable alternative conformation during the unfolding. The alternative conformation is probably very similar in both cases because the contour length increment obtained is the same as it is for the native protein. However, it is clear from the previous measurements that at equilibrium, the presence of the C-terminal domain alternative conformation is never observed (Figs. 5*A* and 6 (*C* and *G*)). Thus, these refolding experiments indicate that both the N-terminal domain and the C-terminal domain may transiently sample the alternative conformation during refolding. These observations suggest that the alternative conformations of the N- and C-terminal domains may possibly be folding intermediates; however, this warrants further investigation.

To better understand the effects of Ca^{2+} on kinetics, we modeled the refolding as a three-state system, where the states

2–5 in the taxonomy represent a partially folded ensemble of conformations that may transit to the fully folded state (number 6 in taxonomy) or transition from the unfolded state (number 1 in taxonomy) we can describe the kinetics by the three-state process (see "Experimental Procedures"). The refolding experiments provide time points for 0.3–10.3 s, and the proportion of recordings calculated to be in the folded state or partially folded from unfolding at equilibrium can be considered to be at a time point of ∞. These data can be fit by a model of kinetics that uses three coupled ordinary differential equations (Fig. 15; see "Experimental Procedures").

Under the simple complete three-state model (detailed under "Experimental Procedures"), the fit kinetics for Protein S without CaCl$_2$ are as follows.

$$U \underset{9 \pm 3 \times 10^{-3}}{\overset{0.5 \pm 0.04}{\rightleftharpoons}} P \underset{8 \pm 6 \times 10^{-3}}{\overset{3 \pm 0.8 \times 10^{-2}}{\rightleftharpoons}} F$$

$$U \underset{8 \pm 7 \times 10^{-3}}{\overset{5 \pm 3 \times 10^{-2}}{\rightleftharpoons}} F$$

Reactions 1 and 2

The fit kinetics for Protein S with CaCl$_2$ are as follows.

$$U \underset{8 \pm 6 \times 10^{-3}}{\overset{0.7 \pm 0.04}{\rightleftharpoons}} P \underset{6 \pm 4 \times 10^{-3}}{\overset{0.1 \pm 0.02}{\rightleftharpoons}} F$$

$$U \underset{2 \pm 3 \times 10^{-4}}{\overset{0.1 \pm 0.05}{\rightleftharpoons}} F$$

Reactions 3 and 4

As described previously, U is unfolded, P is partially folded, and F is fully folded (all rate constants are in units s^{-1}). The errors shown are S.D. from iterations from a randomized seed of the simplex algorithm fitting (see "Experimental Procedures").

These kinetics indicate that the presence of Ca^{2+} increases the P → F transition rate and may decrease the F → U transition rate, which implies that the native states may be stabilized with Ca^{2+}, and the barrier to the folded state may be diminished (Fig. 15). However, we note that because our data include no information about the interconversion between states 2–5, we cannot rule out a higher order model incorporating more states. Also, the unfolding rates determined here imply a waiting time much longer than the time span of our experiment, so these rates should be viewed as upper limits on the unfolding rate. We are further investigating this kinetic scheme using stop-flow measurements and force-clamp spectroscopy.

Configuration of Protein S Alternative Conformation—How do we interpret the nature of the alternative conformations for both the PS N-terminal and C-terminal domains? To obtain a possible interpretation, we used coarse grain simulations, which can simulate topological and energetic factors to compute probable transition states of proteins (49, 50). The protein was modeled as a Cα model (51) and denatured at a high tem-
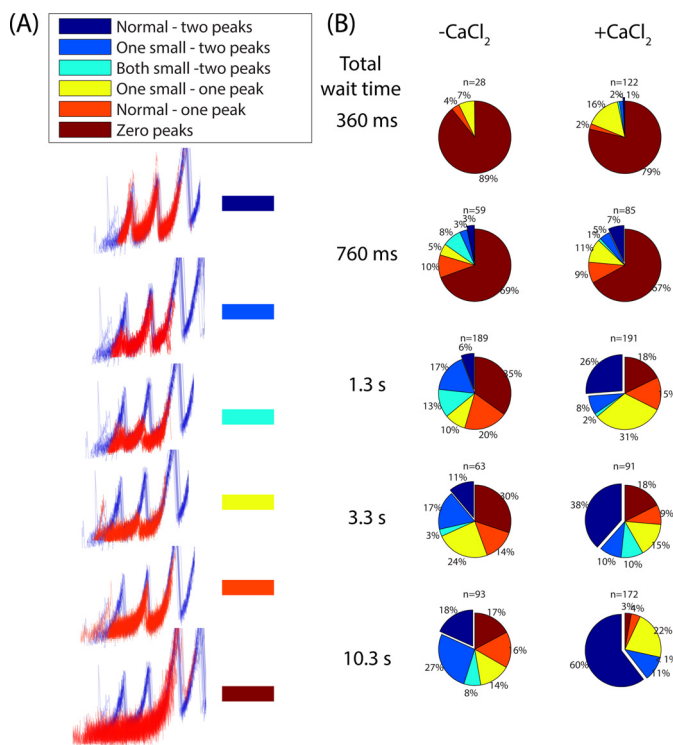

FIGURE 14. **Statistics from the refolding experiments on Protein S.** Six types of events were observed (*A*) with examples in *red* overlaid with the normal unfolding of Protein S (*blue*). The proportions of each type (*B*) are with CaCl$_2$ (*left*) and without CaCl$_2$ (*right*). All proportions have 95% confidence intervals between ±5% and ±10% of the stated value (binomial confidence intervals).
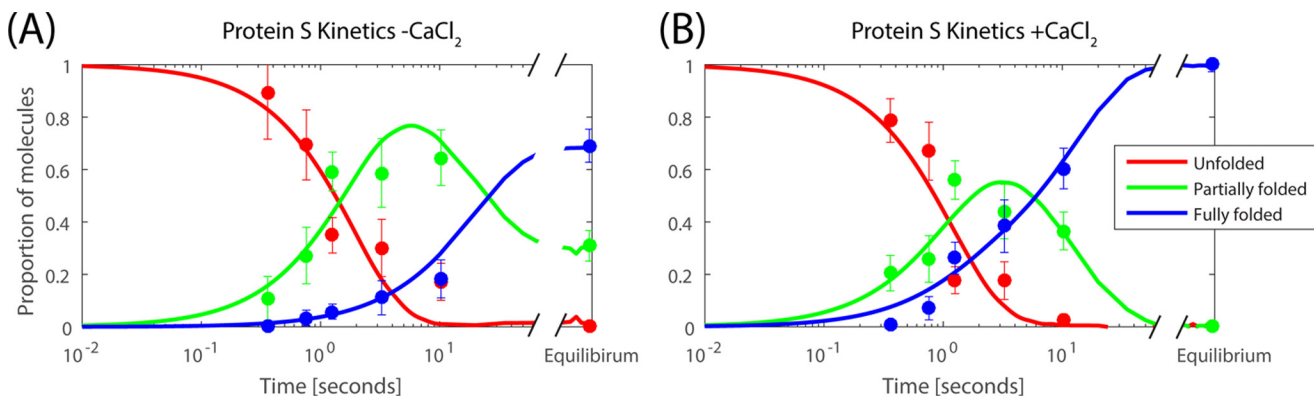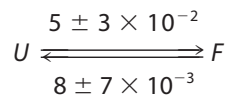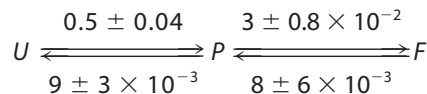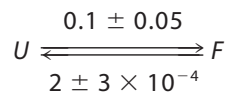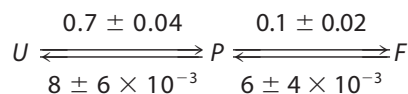

FIGURE 15. *A* and *B*, complete three-state kinetic model fits for Protein S folding with (*A*) and without (*B*) CaCl$_2$. *Solid lines*, model fit; *dots*, experimentally measured proportions of molecules with S.D. Equilibrium measurements come from unfolding results (Fig. 6).
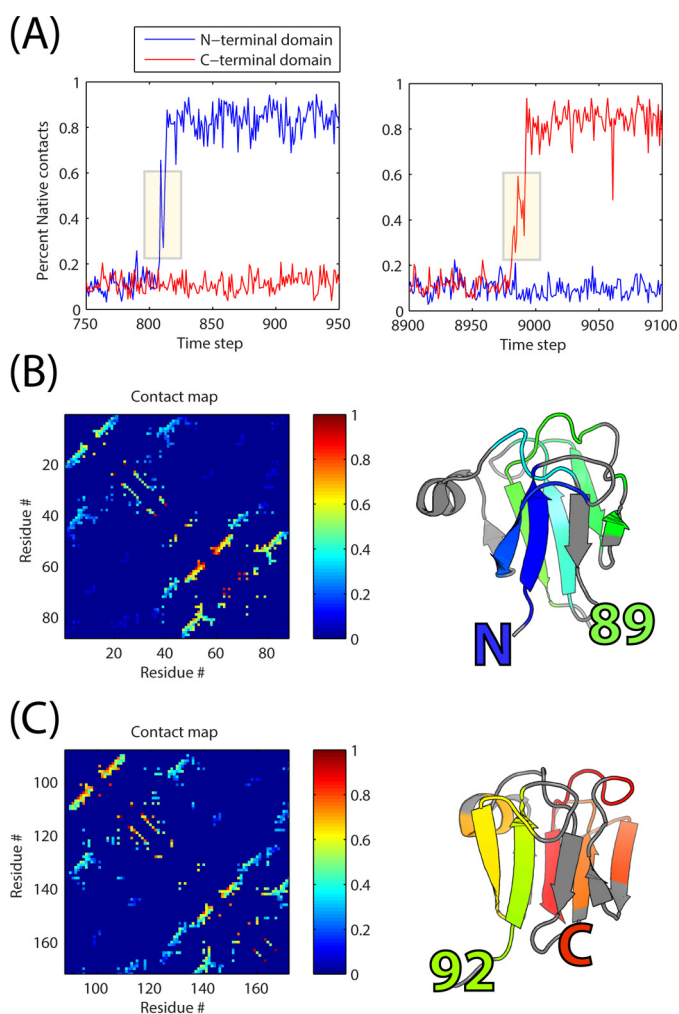
FIGURE 16. **Simulations indicate an interpretation of the alternative conformation.** *A,* two examples of the transient partially folded state obtained when the N-terminal domain folds (*left*) or when the C-terminal domain folds (*right*) where there is a transient state at which 25 and 60% native contacts are formed. The contact maps for the N-terminal intermediate (*B*) and the C-terminal intermediate (*C*) are shown, *colored* by the relative frequency of interaction. The structures are *colored* to show the alternative conformations, where *colored* residues represent secondary structure that is present at least 50% of the time.

perature. The temperature was then quenched to a folding temperature (see "Experimental Procedures"), and the folding was quantified by the percentage of native contacts. Refolding proceeds either from the N-terminal domain first or the C-terminal domain first, but results for the transition states for both were similar, as shown in Fig. 16.

The configurations between the unfolded and folded state (percentage of native contacts between 25 and 60%) showed a distinct configuration consisting of five of the core β-strands in the β Greek key fold. The results were similar for the N-terminal domain and C-terminal domain. In both cases, the β strands on either end of a domain were in the native state, thus clinching the protein into a metastable fold. This structure, if measured using force spectroscopy, would reveal a contour length increment of the entire protein (because the end β-strands are clinched) and would probably rupture at a force that is lower than the normal rupture force of the domain due to the lack of a fully native configuration.

Although charge is not represented in the coarse grain model, it can be seen that the charged residues of the N-terminal domain (residues 10–11 and 70–71) do not form during the transition state (Fig. 16*B*). The fact that these residues only form after the core of the protein folds may indicate why these residues only have a minor influence on the N-terminal domain, to decrease the mechanical stability slightly. Thus, the interpretation of these simulations has qualitative agreement with the force spectroscopy measurements of the alternative conformation and also provides new evidence to support previously hypothesized folding routes of a single domain of Protein S (39).

Discussion

Here we study a calcium-dependent, multidomain protein from the $\beta\gamma$-crystallin family, Protein S, to ascertain, for the first time, the interaction between protein domains and calcium dependence of their folding and unfolding pathways. Using SMFS, we found that Protein S exists in two stable conformations in the absence of Ca^{2+}: a mechanically strong native conformation and a mechanically weaker alternative conformation. Through mutagenesis experiments and computer simulation, we determined that the alternative conformation resides in the N-terminal domain and is caused by structural instability due to intradomain charge repulsion in a high density cluster of negatively charged residues. The alternative conformation becomes undetectable in the presence of Ca^{2+} or acidic pH, because the cluster of negatively charged residues becomes partially or completely neutralized. Interestingly, the equilibrium between the conformations shifts from the alternative conformation to the native conformation when the C-terminal domain is present, because a C-terminal lysine partially neutralizes the negative charge in the N-terminal domain, in the absence of calcium (Fig. 8). These data offer three significant observations and conclusions.

First, unlike in other multidomain proteins, the alternative conformation of Protein S is not a molten globule; rather, it is a highly structured, mechanically stable state. In protein evolution, and in protein design, a cluster of negatively charged residues can coordinate metals, which, generally, will promote stability, by creating new strong ionic bonds between distant residues. However, many proteins that have a cluster of negatively charged residues are found to be unstable in the absence of a metal (31–34, 52–55), probably due to the charge repulsion. Similar to Protein S, this charge-induced instability often can manifest itself in a lower mechanical stability (56). This instability can be alleviated by mutating the electronegative residues to neutral residues, which consequently increases the protein stability in the absence of calcium, such as what we observed in Protein S and what has been observed elsewhere (52). However, the most frequent observation from studies on calcium-binding proteins is that the absence of metal produces a "molten globule" (35, 54, 55, 57). Our data show that, instead of a molten globule, the alternative conformation is defined by a consistent mechanical stability, and the contour length increment following its mechanical denaturation remains the same as for the native ensemble (Fig. 3). It is tempting to speculate that the molten globule structures observed in bulk measurements, in
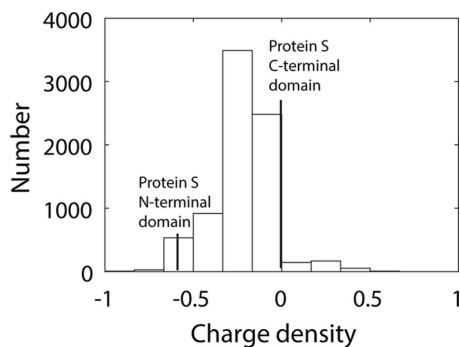
FIGURE 17. **The charge density of each binding site and flanking residues of all known calcium-binding proteins is plotted (for more details, see "Experimental Procedures").** The charge density of the binding site in the N-terminal domain of Protein S (-0.66) is higher than in 99% of all known proteins. In contrast, the charge density of the C-terminal domain is neutral.

the absence of metal ligand, may actually preserve some residual structures that could be identified through SMFS measurements.

Second, we discovered a compensatory mechanism in Protein S that provides stability in the absence of calcium. This compensatory mechanism seems to be only available to multidomain or multimeric proteins; the C-terminal domain of Protein S has a lysine that binds to the charged pocket of the N-terminal domain (Fig. 8). Simulations showed that this lysine prevents destabilization of the N-terminal loop (Fig. 10), whereas experiments show that the C-terminal domain helps to shift the ensemble from 55% in the alternative conformation (Fig. 3) to only 34% in the alternative conformation (Fig. 6) in the absence of calcium. This mechanism is a novel property that multidomain proteins may have evolved to provide another level of stability control, especially in the absence of calcium.

Third, there are many multidomain proteins that have exceptionally electronegative binding sites, similar to Protein S. These binding sites are designed to allow proteins to buffer calcium levels (5) or for producing conformational changes that can stimulate cellular signaling (3). An analysis of all known calcium-binding proteins (see "Experimental Procedures") shows that only 1% of all known calcium-binding proteins have sites with charge density similar to or higher than the site on the N-terminal domain of Protein S (Fig. 17; a full list is included in supplemental Table 1). These proteins with highly charged binding sites share common biological functions for conformational switching or calcium titration; for example, calsequestrin (58), human Hsp70 (59), and an Na^+/Ca^{2+} exchanger binding domain (60) have a similarly high charge density binding pocket as Protein S. Many of these proteins with exceptionally electronegative binding sites have already been reported to harbor great structural instability in the absence of calcium (54, 55, 57–62) or have increases in folding speed due to calcium (63). Because many of these proteins are also multidomain, our results on Protein S may apply to many of the proteins with high charge density sites. Calcium binding is one of the most important and ubiquitous mechanisms for signaling, which is maintained through a variety of proteins, especially multidomain proteins.

Through Protein S we have shown that there exists a multitude of ways to modulate domain stability through highly charged calcium-binding sites and domain interactions. Because the majority of proteins are multidomain, we speculate that these mechanisms of interaction may be highly prevalent and help to avoid aggregation-prone molten globules and instead allow well defined conformational ensembles in the unbound state.

Experimental Procedures

Protein Engineering—All proteins for force spectroscopy experiments (Protein S, Protein S D11Q, E70P, and E71T, truncated N-terminal domain, truncated C-terminal domain) were designed to have I27 flanking the protein of interest. The addition of flanking I27 domains serves to facilitate the pickup of the molecule and also serves as a positive control and fingerprint of a single-molecule event (64). All Protein S sequences were synthesized (GenScript USA Inc., Piscataway, NJ) and placed into the third module of the poly(I27) pRSETa vector, a kind gift from Jane Clarke (65). The eighth module was replaced with a Strep-Tag. All engineered plasmids were transformed into *Escherichia coli* C41(DE3)pLysS cells (Lucigen Corp., Middleton, WI), and expression was induced using isopropyl 1-thio-β-D-galactopyranoside. Proteins were purified by running cell lysate through a Strep-Tag-Strep-Tactin column (IBA, Olivette, MO). Protein was then run through size exclusion columns and resuspended in Tris-HCl, pH 7.2, and stored at 4 °C for subsequent use. Fresh protein was prepared every 2 weeks for experiments.

AFM Spectroscopy—All AFM measurements were obtained using a custom-built AFM instrument (66). Automation routines to control the AFM (67) were implemented in Labview version 7.0 (National Instruments, Austin, TX). Cantilever spring constants were calibrated in the buffer solution using the energy equipartition theorem (68). Purified protein was dialyzed and concentrated in Tris-HCl, pH 7.2, with 1 mM EGTA for experiments without Ca^{2+} or dialyzed into Tris-HCl, pH 7.2, with 4 mM $CaCl_2$ for experiments with Ca^{2+}. Measurements were done in either Tris-HCl, pH 7.2, with 1 mM EDTA; Tris-HCl, pH 7.2, with 2 mM $CaCl_2$; or citrate buffer, pH 3.7–4.5 (when specified). In all experiments, the purified protein was diluted to 50–150 μg/ml.

Protein attachment to the substrate and cantilever was achieved using nonspecific absorption, without any tip functionalization. Slides for AFM spectroscopy were prepared by first incubating 100 μl of diluted protein onto a gold slide, for 1 h. Gold slides were prepared by cleaning glass with Piranha solution and then evaporating 80 nm of chromium and 250 nm of gold onto the surface. Gold slides were stored in argon to prevent oxidation until use. Protein on the gold surface stuck nonspecifically through absorption. To remove unadsorbed protein, we carefully drew off 60 μl from the deposited protein on the slide before measurements.
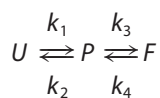
Measurements were performed using MLCT cantilevers (Bruker, Camarillo, CA), which have a spring constant of 51 $\pm$ 10 pN/nm. Each force-extension profile was generated by a constant velocity extension and retraction at a speed of 300 nm/s. Protein was adsorbed to the cantilever nonspecifically, by

bringing the cantilever in contact with the surface with a force of ~200 pN. If a protein molecule adsorbed to the cantilever tip and also to the surface, the retraction of the cantilever would unravel the protein through the tension applied during the measurement. A single-molecule recording was determined through the presence of at least four I27 domains, which would guarantee that the middle domain has been unfolded. A worm-like chain model (24) with persistence length of 0.4 nm was fit to each peak in order to measure contour length increments in the force-extension data (69).

Coarse Grain Simulation—Structure-based models were generated using the SMOG Web server version 1.2.2 (51) from Protein Data Bank entry 1PRS (70). In this coarse grain model, each residue is modeled as a single sphere, and forces are determined by a coarse grain potential. This potential contains terms for bonds, angles, and improper angles, which have equilibrium values based on the initial structure. All residues identified as a contact have an attractive 12-6 potential, and residues identified as non-contacts have a repulsive 12 potential. For more information about parameter values, see Clementi *et al.* (49). The temperature used for all calculations was the folding temperature at which the folded state and the unfolded state are equally populated. The folding temperature was found by determining where the specific heat is maximal. Simulations were conducted using GROMACS version 4.5.5 (71).

All-atom Molecular Dynamics Simulation—All-atom models for explicit solvent molecular dynamics were generated from Protein Data Bank entry 1PRS (70). The truncated N-terminal domain was generated by deleting the last 83 residues, and the truncated C-terminal domain was generated by deleting the first 89 residues. Hydrogen atoms were added using VMD Automatic PSF Builder (72). A water box was created around each model to encompass the structure and allow a 2.8-nm water margin from each side. The salt was added to replace water molecules until the concentration reached 150 mM, and the charge was neutralized. For simulations with $CaCl_2$, the bound Ca^{2+} in the 1PRS structure was not removed, and additional Ca^{2+} and Cl^- ions were added to a concentration of 1 mM. Structures were then minimized and equilibrated at constant pressure for 1 ns using NAMD2 version 2.10b1 (73). Production simulations were then done for 10 ns at constant volume.

Modeling Refolding Kinetics—Kinetics were modeled using the complete three-state folding model of a protein,

$$U \underset{k_2}{\overset{k_1}{\rightleftharpoons}} P \underset{k_4}{\overset{k_3}{\rightleftharpoons}} F$$

Reaction 5

and

$$U \underset{k_6}{\overset{k_5}{\rightleftharpoons}} F$$

Reaction 6

which is described by the three ordinary differential equations,

$$\frac{dU(t)}{dt} = -(k_1 + k_5)U(t) + k_2P(t) + k_6F(t) \qquad \text{(Eq. 1)}$$

$$\frac{dP(t)}{dt} = k_1U(t) - k_2P(t) - k_3P(t) + k_4F(t) \qquad \text{(Eq. 2)}$$

$$\frac{dF(t)}{dt} = k_5U(t) + k_3P(t) - (k_4 + k_6)F(t) \qquad \text{(Eq. 3)}$$

where $U(t)$, $P(t)$, and $F(t)$ are the normalized probability of detecting the unfolded, partially folded, or fully folded state at time t, respectively. The initial conditions were such that $(U,P,F) = (1,0,0)$ at $t = 0$. The final conditions come from unfolding, where (U,P,F) reflects the proportion of unfolded, partially folded, or fully folded molecules determined from fully unfolding molecules. Partially folded events were classified by having an unfolding force less than two S.D. values from the mean of the unfolding distribution (or the higher mean of the two unfolding distributions in the case of multimodal distributions). The data for the intermediate time points come from refolding experiments using the same classification criterion. The 95% confidence intervals for these proportions were determined using binomial proportion confidence intervals. The three-state folding model equations are solved numerically using a simplex solver that minimizes the difference between the experimentally measured and calculated $U(t)$, $P(t)$, and $F(t)$. The simplex solver was run for 500 iterations, although it converged after 200 iterations. The errors estimated for the rates are from 20 iterations from the simplex solver starting from randomized initial conditions.

Analysis of Known Protein Charge Densities—We analyzed the charge density by computing the total charge of a binding site and dividing by the number of residues in the binding site. A binding site was defined to include all of the known binding residues *and* the immediate flanking residues. Protein sequence and feature information was downloaded from the NCBI Protein database using the query "het[Feature key] OR Het[Feature key] OR HET[Feature key] OR Ca2[Title] OR Calcium[Title] OR ca2[Title] OR calcium[Title]". The resulting 79,297 data files were parsed for heterogens of the type "CA" and their binding sites were located. The data were analyzed using custom Python scripts utilizing Biopython (74) to parse GenPept data. Data deduplication was performed by merging entries that contained the same amino acid sequence and the same binding sites. Entries that had only one amino acid in the binding were not used for analysis; this procedure resulted in 7,830 entries (for the list, see supplemental Table 1). Each of these entries had the total charge determined using the sequence and the binding site information, computing −1 for D and E, +1 for R and K, and 0 otherwise. The results for the charge density of each binding site in each protein are plotted in Fig. 17.

Alternative Conformations of Protein S

References

1. Braselmann, E., Chaney, J. L., and Clark, P. L. (2013) Folding the proteome. *Trends Biochem. Sci.* **38,** 337–344
2. Tompa, P. (2012) On the supertertiary structure of proteins. *Nat. Chem. Biol.* **8,** 597–600
3. Clapham, D. E. (1995) Calcium signaling. *Cell* **80,** 259–268
4. Berridge, M. J., Bootman, M. D., and Roderick, H. L. (2003) Calcium signalling: dynamics, homeostasis and remodelling. *Nat. Rev. Mol. Cell Biol.* **4,** 517–529
5. Carafoli, E. (1987) Intracellular calcium homeostasis. *Annu. Rev. Biochem.* **56,** 395–433
6. Ray, M. E., Wistow, G., Su, Y. A., Meltzer, P. S., and Trent, J. M. (1997) AIM1, a novel non-lens member of the $\beta\gamma$-crystallin superfamily, is associated with the control of tumorigenicity in human malignant melanoma. *Proc. Natl. Acad. Sci. U.S.A.* **94,** 3229–3234
7. Sun, H., Ma, Z., Li, Y., Liu, B., Li, Z., Ding, X., Gao, Y., Ma, W., Tang, X., Li, X., and Shen, Y. (2005) γ-s crystallin gene (crygs) mutation causes dominant progressive cortical cataract in humans. *J. Med. Genet.* **42,** 706–710
8. Ma, Z., Piszczek, G., Wingfield, P. T., Sergeev, Y. V., and Hejtmancik, J. F. (2009) The g18v crygs mutation associated with human cataracts increases γS-crystallin sensitivity to thermal and chemical stress. *Biochemistry* **48,** 7334–7341
9. Stigler, J., Ziegler, F., Gieseke, A., Gebhardt, J. C. M., and Rief, M. (2011) The complex folding network of single calmodulin molecules. *Science* **334,** 512–516
10. Garcia-Manyes, S., Giganti, D., Badilla, C. L., Lezamiz, A., Perales-Calvo, J., Beedle, A. E., and Fernández, J. M. (2016) Single molecule force spectroscopy predicts a misfolded, domain-swapped conformation in human γD-crystallin. *J. Biol. Chem.* **291,** 4226–4235
11. Neuman, K. C., and Nagy, A. (2008) Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy. *Nat. Methods* **5,** 491–505
12. Borgia, A., Williams, P. M., and Clarke, J. (2008) Single-molecule studies of protein folding. *Annu. Rev. Biochem.* **77,** 101–125
13. Hoffmann, T., and Dougan, L. (2012) Single molecule force spectroscopy using polyproteins. *Chem. Soc. Rev.* **41,** 4781–4796
14. Noy, A., and Friddle, R. W. (2013) Practical single molecule force spectroscopy: how to determine fundamental thermodynamic parameters of intermolecular bonds with an atomic force microscope. *Methods* **60,** 142–150
15. Žoldák, G., and Rief, M. (2013) Force as a single molecule probe of multidimensional protein energy landscapes. *Curr. Opin. Struct. Biol.* **23,** 48–57
16. Woodside, M. T., and Block, S. M. (2014) Reconstructing folding energy landscapes by single-molecule force spectroscopy. *Annu. Rev. Biophys.* **43,** 19–39
17. Fitter, J. (2009) The perspectives of studying multi-domain protein folding. *Cell. Mol. Life Sci.* **66,** 1672–1681
18. Scholl, Z. N., Yang, W., and Marszalek, P. E. (2014) Chaperones rescue luciferase folding by separating its domains. *J. Biol. Chem.* **289,** 28607–28618
19. Kessler, M., and Gaub, H. E. (2006) Unfolding barriers in bacteriorhodopsin probed from the cytoplasmic and the extracellular side by afm. *Structure* **14,** 521–527
20. Lee, W., Zeng, X., Zhou, H.-X., Bennett, V., Yang, W., and Marszalek, P. E. (2010) Full reconstruction of a vectorial protein folding pathway by atomic force microscopy and molecular dynamics simulations. *J. Biol. Chem.* **285,** 38167–38172
21. Han, J.-H., Batey, S., Nickson, A. A., Teichmann, S. A., and Clarke, J. (2007) The folding and evolution of multidomain proteins. *Nat. Rev. Mol. Cell Biol.* **8,** 319–330
22. Aravind, P., Mishra, A., Suman, S. K., Jobby, M. K., Sankaranarayanan, R., and Sharma, Y. (2009) The $\beta\gamma$-crystallin superfamily contains a universal motif for binding calcium. *Biochemistry* **48,** 12180–12190
23. Wenk, M., Jaenicke, R., and Mayr, E.-M. (1998) Kinetic stabilisation of a modular protein by domain interactions. *FEBS Lett.* **438,** 127–130
24. Wistow, G. (1990) Evolution of a protein superfamily: relationships between vertebrate lens crystallins and microorganism dormancy proteins. *J. Mol. Evol.* **30,** 140–145
25. Aravind, P., Wistow, G., Sharma, Y., and Sankaranarayanan, R. (2008) Exploring the limits of sequence and structure in a variant $\beta\gamma$-crystallin domain of the protein absent in melanoma-1 (aim1). *J. Mol. Biol.* **381,** 509–518
26. Clout, N. J., Kretschmar, M., Jaenicke, R., and Slingsby, C. (2001) Crystal structure of the calcium-loaded spherulin 3a dimer sheds light on the evolution of the eye lens $\beta\gamma$-crystallin domain fold. *Structure* **9,** 115–124
27. Lubsen, N. H., Aarts, H. J., and Schoenmakers, J. G. (1988) The evolution of lenticular proteins: the β- and γ-crystallin super gene family. *Prog. Biophys. Mol. Biol.* **51,** 47–76
28. Sharma, Y., Rao, C. M., Narasu, M. L., Rao, S. C., Somasundaram, T., Gopalakrishna, A., and Balasubramanian, D. (1989) Calcium ion binding to δ- and to β-crystallins: the presence of the "EF-hand" motif in δ-crystallin that aids in calcium ion binding. *J. Biol. Chem.* **264,** 12794–12799
29. Rajini, B., Shridas, P., Sundari, C. S., Muralidhar, D., Chandani, S., Thomas, F., and Sharma, Y. (2001) Calcium binding properties of γ-crystallin calcium ion binds at the Greek key $\beta\gamma$-crystallin fold. *J. Biol. Chem.* **276,** 38464–38471
30. Srivastava, S. S., Mishra, A., Krishnan, B., and Sharma, Y. (2014) Ca^{2+}-binding motif of $\beta\gamma$-crystallins. *J. Biol. Chem.* **289,** 10958–10966
31. Kretschmar, M., Mayr, E.-M., and Jaenicke, R. (1999) Kinetic and thermodynamic stabilization of the $\beta\gamma$-crystallin homolog spherulin 3a from physarum polycephalum by calcium binding. *J. Mol. Biol.* **289,** 701–705
32. Giancola, C., Pizzo, E., Di Maro, A., Cubellis, M. V., and D'Alessio, G. (2005) Preparation and characterization of geodin. *FEBS J.* **272,** 1023–1035
33. Suman, S. K., Mishra, A., Ravindra, D., Yeramala, L., and Sharma, Y. (2011) Evolutionary remodeling of $\beta\gamma$-crystallins for domain stability at Cost of ca^{2+} binding. *J. Biol. Chem.* **286,** 43891–43901
34. Barnwal, R. P., Jobby, M. K., Devi, K. M., Sharma, Y., and Chary, K. V. (2009) Solution structure and calcium-binding properties of m-crystallin, a primordial $\beta\gamma$-crystallin from archaea. *J. Mol. Biol.* **386,** 675–689
35. Jobby, M. K., and Sharma, Y. (2005) Calcium-binding crystallins from *Yersinia pestis*: characterization of two single $\beta\gamma$-crystallin domains of a putative exported protein. *J. Biol. Chem.* **280,** 1209–1216
36. Slingsby, C., Wistow, G. J., and Clark, A. R. (2013) Evolution of crystallins for a role in the vertebrate eye lens. *Protein Sci.* **22,** 367–380
37. Wenk, M., Baumgartner, R., Holak, T. A., Huber, R., Jaenicke, R., and Mayr, E.-M. (1999) The domains of protein S from *Myxococcus xanthus*: structure, stability and interactions. *J. Mol. Biol.* **286,** 1533–1545
38. Wenk, M., and Jaenicke, R. (1999) Calorimetric analysis of the Ca^{2+}-binding $\beta\gamma$-crystallin homolog protein S from *Myxococcus xanthus*: intrinsic stability and mutual stabilization of domains. *J. Mol. Biol.* **293,** 117–124
39. Bagby, S., Go, S., Inouye, S., Ikura, M., and Chakrabartty, A. (1998) Equilibrium folding intermediates of a Greek key β-barrel protein. *J. Mol. Biol.* **276,** 669–681
40. Rief, M., Gautel, M., Oesterhelt, F., Fernandez, J. M., and Gaub, H. E. (1997) Reversible unfolding of individual titin immunoglobulin domains by afm. *Science* **276,** 1109–1112
41. Carrion-Vazquez, M., Oberhauser, A. F., Fowler, S. B., Marszalek, P. E., Broedel, S. E., Clarke, J., and Fernandez, J. M. (1999) Mechanical and chemical unfolding of a single protein: a comparison. *Proc. Natl. Acad. Sci. U.S.A.* **96,** 3694–3699
42. Dietz, H., and Rief, M. (2006) Protein structure by mechanical triangulation. *Proc. Natl. Acad. Sci. U.S.A.* **103,** 1244–1247
43. Garai, A., Zhang, Y., and Dudko, O. K. (2014) Conformational dynamics through an intermediate. *J. Chem. Phys.* **140,** 135101
44. Scholl, Z. N., Yang, W., and Marszalek, P. E. (2015) Direct observation of multimer stabilization in the mechanical unfolding pathway of a protein undergoing oligomerization. *ACS Nano* **9,** 1189–1197

45. Kotamarthi, H. C., Sharma, R., and Koti Ainavarapu, S. R. (2013) Single-molecule studies on PolySUMO proteins reveal their mechanical flexibility. *Biophys. J.* **104,** 2273–2281

46. Teintze, M., Inouye, M., and Inouye, S. (1988) Characterization of calcium-binding sites in development-specific protein S of *Myxococcus xanthus* using site-specific mutagenesis. *J. Biol. Chem.* **263,** 1199–1203

47. Andreotti, A. H. (2003) Native state proline isomerization: an intrinsic molecular switch. *Biochemistry* **42,** 9515–9524

48. Oberhauser, A. F., Marszalek, P. E., Erickson, H. P., and Fernandez, J. M. (1998) The molecular elasticity of the extracellular matrix protein tenascin. *Nature* **393,** 181–185

49. Clementi, C., Nymeyer, H., and Onuchic, J. N. (2000) Topological and energetic factors: what determines the structural details of the transition state ensemble and "en-route" intermediates for protein folding? an investigation for small globular proteins. *J. Mol. Biol.* **298,** 937–953

50. Clementi, C., Jennings, P. A., and Onuchic, J. N. (2001) Prediction of folding mechanism for circular-permuted proteins. *J. Mol. Biol.* **311,** 879–890

51. Noel, J. K., Whitford, P. C., Sanbonmatsu, K. Y., and Onuchic, J. N. (2010) SMOG@ ctbp: simplified deployment of structure-based models in gromacs. *Nucleic Acids Res.* **38,** W657–W661

52. Spurway, T. D., Morland, C., Cooper, A., Sumner, I., Hazlewood, G. P., O'Donnell, A. G., Pickersgill, R. W., and Gilbert, H. J. (1997) Calcium protects a mesophilic xylanase from proteinase inactivation and thermal unfolding. *J. Biol. Chem.* **272,** 17523–17530

53. Yamniuk, A. P., Silver, D. M., Anderson, K. L., Martin, S. R., and Vogel, H. J. (2007) Domain stability and metal-induced folding of calcium-and integrin-binding protein 1. *Biochemistry* **46,** 7088–7098

54. Yamniuk, A. P., Nguyen, L. T., Hoang, T. T., and Vogel, H. J. (2004) Metal ion binding properties and conformational states of calcium-and integrin-binding protein. *Biochemistry* **43,** 2558–2568

55. Takeuchi, Y., Tanaka, S., Matsumura, H., Koga, Y., Takano, K., and Kanaya, S. (2009) Requirement of a unique Ca^{2+}-binding loop for folding of Tk-subtilisin from a hyperthermophilic archaeon. *Biochemistry* **48,** 10637–10643

56. Cao, Y., Yoo, T., and Li, H. (2008) Single molecule force spectroscopy reveals engineered metal chelation is a general approach to enhance mechanical stability of proteins. *Proc. Natl. Acad. Sci. U.S.A.* **105,** 11152–11157

57. Permyakov, E. A., and Berliner, L. J. (2000) α-Lactalbumin: structure and function. *FEBS Lett.* **473,** 269–274

58. Lewis, K. M., Ronish, L. A., Ríos, E., and Kang, C. (2015) Characterization of two human skeletal calsequestrin mutants implicated in malignant hyperthermia and vacuolar aggregate myopathy. *J. Biol. Chem.* **290,** 28665–28674

59. Sriram, M., Osipiuk, J., Freeman, B., Morimoto, R., and Joachimiak, A. (1997) Human hsp70 molecular chaperone binds two calcium ions within the atpase domain. *Structure* **5,** 403–414

60. Hilge, M., Aelen, J., and Vuister, G. W. (2006) Ca^{2+} regulation in the Na^+/Ca^{2+} exchanger involves two markedly different Ca^{2+} sensors. *Mol. Cell.* **22,** 15–25

61. Sotomayor-Pérez, A. C., Ladant, D., and Chenal, A. (2011) Calcium-induced folding of intrinsically disordered repeat-in-toxin (rtx) motifs via changes of protein charges and oligomerization states. *J. Biol. Chem.* **286,** 16997–17004

62. Lowry, C. L., McGeehan, G., and LeVine, H., 3rd (1992) Metal ion stabilization of the conformation of a recombinant 19-kDa catalytic fragment of human fibroblast collagenase. *Proteins* **12,** 42–48

63. Xu, A. J., and Springer, T. A. (2012) Calcium stabilizes the von Willebrand factor a2 domain by promoting refolding. *Proc. Natl. Acad. Sci. U.S.A.* **109,** 3742–3747

64. Scholl, Z. N., Li, Q., and Marszalek, P. E. (2014) Single molecule mechanical manipulation for studying biological properties of proteins, DNA, and sugars. *Wiley Interdiscip. Rev. Nanomed. Nanobiotechnol.* **6,** 211–229

65. Steward, A., Toca-Herrera, J. L., and Clarke, J. (2002) Versatile cloning system for construction of multimeric proteins for use in atomic force microscopy. *Protein Sci.* **11,** 2179–2183

66. Lee, G., Abdi, K., Jiang, Y., Michaely, P., Bennett, V., and Marszalek, P. E. (2006) Nanospring behaviour of ankyrin repeats. *Nature* **440,** 246–249

67. Scholl, Z. N., and Marszalek, P. E. (2014) Improving single molecule force spectroscopy through automated real-time data collection and quantification of experimental conditions. *Ultramicroscopy* **136,** 7–14

68. Florin, E. L., Rief, M., Lehmann, H., Ludwig, M., Dornmair, C., Moy, V. T., and Gaub, H. E. (1995) Sensing specific molecular-interactions with the atomic-force microscope. *Biosensors Bioelectronics* **10,** 895–901

69. Marko, J. F., and Siggia, E. D. (1995) Stretching DNA. *Macromolecules* **28,** 8759–8770

70. Bagby, S., Harvey, T. S., Eagle, S. G., Inouye, S., and Ikura, M. (1994) NMR-derived three-dimensional solution structure of protein S complexed with calcium. *Structure* **2,** 107–122

71. Pronk, S., Páll, S., Schulz, R., Larsson, P., Bjelkmar, P., Apostolov, R., Shirts, M. R., Smith, J. C., Kasson, P. M., van der Spoel, D., Hess, B., and Lindahl, E. (2013) GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* **29,** 845–854

72. Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: visual molecular dynamics. *J. Mol. Graphics* **14,** 33–38, 27–28

73. Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kalé, L., and Schulten, K. (2005) Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26,** 1781–1802

74. Cock, P. J., Antao, T., Chang, J. T., Chapman, B. A., Cox, C. J., Dalke, A., Friedberg, I., Hamelryck, T., Kauff, F., Wilczynski, B., and de Hoon, M. J. (2009) Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics* **25,** 1422–1423