Chaperones Rescue Luciferase Folding by Separating Its Domains*

Over the last 50 years, significant progress has been made toward understanding how small single-domain proteins fold. However, very little is known about folding mechanisms of medium and large multidomain proteins that predominate the proteomes of all forms of life. Large proteins frequently fold cotranslationally and/or require chaperones. Firefly (Photinus pyralis) luciferase (Luciferase, 550 residues) has been a model of a cotranslationally folding protein whose extremely slow refolding (approximately days) is catalyzed by chaperones. However, the mechanism by which Luciferase misfolds and how chaperones assist Luciferase refolding remains unknown. Here we combine single-molecule force spectroscopy (atomic force microscopy (AFM)/single-molecule force spectroscopy) with steered molecular dynamic computer simulations to unravel the mechanism of chaperone-assisted Luciferase refolding. Our AFM and steered molecular dynamic results show that partially unfolded Luciferase, with the N-terminal domain remaining folded, can refold robustly without chaperones. Complete unfolding causes Luciferase to get trapped in very stable non-native configurations involving interactions between N- and C-terminal residues. However, chaperones allow the completely unfolded Luciferase to refold quickly in AFM experiments, strongly suggesting that chaperones are able to sequester non-natively contacting residues. More generally, we suggest that many chaperones, rather than actively promoting the folding, mimic the ribosomal exit tunnel and physically separate protein domains, allowing them to fold in a cotranslational-like sequential process.

The structure and function of proteins is fundamental to almost all biological processes, and the understanding of how proteins obtain their structure through folding is crucial for the prevention of many diseases (1, 2). There have been many experiments over 50 years of research on proteins (for example, see reviews in Refs. 3–7), but there are still crucial aspects of protein folding that need further study, one being the understanding of the protein intermediates and the energy landscapes of protein folding (e.g. the space of protein conformations and their associated free energies) (8). A complete understanding of protein folding pathways could allow drugs to be designed to intermediate states in the folding pathway, thus enabling a novel area of therapeutics. Numerous experiments have shown that proteins evolved by smoothing their folding energy landscapes to eliminate low-energy, non-native intermediate conformations (or kinetic traps) (9, 10). However, much of the work on protein folding has been on small single-domain proteins although they only make up less than 30% of the proteomes in all kingdoms of life (11). It has become clear that a simple smooth energy landscape does not apply to multidomain proteins because multidomain proteins can have non-Anfinsen mechanisms like kinetic partitioning (12), cotranslational folding (13, 14), co-existing intersecting folding pathways (15), or confinement by chaperonins (16, 17). Here we study the folding mechanism of firefly (Photinus pyralis) luciferase (Luciferase) as a model of a large multidomain protein and determine whether there are important differences between the folding of multidomain proteins and single-domain proteins.

Luciferase is a monomeric 61-kDa protein that uses MgATP to catalyze the oxidation of D-(-)-Luciferin (Luciferin) to produce a photon (18–20). The structural conformation of the firefly luciferase protein closely resembles other AMP-forming ligases like acyl-CoA ligases (21) and nonribosomal peptide synthetases (22) because it has a large N-terminal domain and a smaller C-terminal domain, and both are crucial for catalysis (23). Previous studies using traditional biochemical methods determined that once unfolded Luciferase does not refold in an appreciable amount of time (∼10–72 h) (24, 25). The folding time of Luciferase decreases by orders of magnitude (to approximately minutes) by folding cotranslationally on the ribosome.

*This work was supported by National Science Foundation Graduate Research Fellowship Program Grant 1106401 (to Z. N. S.) and by NSF Grant MCB-1052208 (to P. E. M. and W. Y.).

1 Supported by the Kimberly-Clark Corporation. To whom correspondence should be addressed: Dept. of Mechanical Engineering and Materials Science, Center for Biologically Inspired Materials and Material Systems, Duke University, Durham, NC 27705. Tel.: 919-660-5381; E-mail: pemar@duke.edu.
Luciferase Refolding Rescued by Chaperones

FIGURE 1. Luciferase activity in the I27-flanked construct, I27⌬Luciferase-I27, was tested for native enzymatic activity. A, B, and D, as shown in C, the protein construct still maintains native activity when cofactors are present, indicating that the protein construct maintains the native structure of the enzyme.

or when given chaperones (24, 26–28). Although these experiments are elucidating, important questions remain: what prevents Luciferase from refolding quickly? How does cotranslational folding prevent misfolding? How exactly do chaperones allow relatively fast refolding? Here we address these questions using single-molecule force spectroscopy (SMFS).2 SMFS has two unique advantages over traditional methods for studying large multidomain proteins. Their unfolding and refolding can be examined for individual proteins in isolation, therefore minimizing the opportunity for their aggregation in the unfolded state, and SMFS allows the precise manipulation of proteins and their subunits and, therefore, offers unprecedented control for denaturing single domains/subunits within multidomain proteins (29, 30). This approach reduces the complexity of the folding reaction and facilitates relating the experimental data to the structural information (29, 31, 32).

EXPERIMENTAL PROCEDURES

Protein Purification—We developed a construct containing full-length firefly luciferase (Promega) or truncated Luciferase (Luc203, containing only residues 1–203) flanked by three I27 domains and four I27 domains at the DNA level (Fig. 1A) and containing an N-terminal His6 tag and a C-terminal Cys2 tag. We purified protein produced from C41(DE3)pLysS cells (Lucigen Corp.) using nickel-nitrilotriacetic acid columns (Qiagen). The resulting purified protein was dialyzed into a buffer containing either 1× or 2× PBS (pH 7.6) at a concentration of 1–10 mg/ml and was used for AFM measurements. The protein was still active when flanked by I27 domains (Fig. 1).

Atomic Force Microscopy—AFM experiments were performed using custom-built instruments (33). Proteins were diluted to 150 μg/ml in a 1× PBS buffer or 2× PBS buffer (for Luc203) and loaded onto a freshly evaporated gold substrate for 1 h. Buffers with ATP also contained 15 mM MgCl2 (to ensure 99% bound), 4 mM EGTA, and 180 mM to 1 mM Luciferin (when noted). For measurements with chaperones, we used PBS with 15 mM MgCl2, 2 mM ATP, and rabbit reticulocyte lysate (untreated) (Promega, catalog no. L4151). The substrate was then washed once and used for pulling experiments. We used either OBL cantilevers (spring constant, ~6 pN/nm) or MLCT cantilevers (spring constant, ~16 pN/nm) (Bruker). During experiments, the AFM cantilever was pressed against the sample at a contact force of 100–500 pN to nonspecifically bind to the protein. The presence of at least five I27 domains (ΔL, ~28 nm; unfolding force, ~200 pN) allowed for the unequivocal determination of the unfolding pattern of the Luciferase protein. The data were analyzed using Matlab 7.10 (Mathworks).

Force extension curves were fit using the worm-like chain (WLC) model. The persistence length for the WLC model was determined in an unbiased way using a global fitting procedure. All curves (>1000) were fit simultaneously with a WLC of various persistence lengths. The root mean square error was calculated for each fit. The root mean square error distribution closely follows the y distribution, so the mode (peak of distribution) was calculated for each persistence length from fitting to a y distribution. The root mean square error mode was plotted against a range of persistence lengths (0.1–1.0 nm), which created a smooth curve with a single minimum. The minimum indicates the best overall fit to all of the data (lowest root mean square error), and the value of the persistence length at this point was used for all subsequent WLC fits: 0.29 nm for I27 and 0.67 nm for peaks of Luciferase.

Determination of Domain Boundaries for Multidomain Proteins from Force Spectroscopy Measurements—Determining the number of amino acids unfolding from a force rip depends on the crystal structure of the corresponding protein. For a simple protein like I27, the number of amino acids is simply equal to the sum of the initial N-C extension from the crystal structure (the “hidden” residues) and the contour length increment measured by the AFM divided by the length per residue. And, vice versa, the length per residue can be calculated if the number of residues is known. Using I27, we calculated the length per residue. The median contour length increment for I27 (28.8 nm) and measured N-C extension (4.2 nm) and number of residues (89) estimated the length per extended residue to be 0.37 nm, which is consistent with the value provided by Dietz and Rief (34).

For a protein composed of intricately interacting domains, however, the simple initial N-C extension is not sufficient to capture the correct number of hidden residues. During the unfolding of each domain and its separation from the rest of the folded structure of the multidomain protein there is an associated reorientation of the remaining structure to align the rest of the protein to the new pulling direction. The realignment can contribute to the experimentally measured contour length increment. Therefore, the contour length increment measured is actually the total length for the residues in the domain minus the difference between the initial length of the new reoriented structure and the previous structure. This effect is illustrated in Fig. 2 and, if neglected, it could significantly increase the uncertainty of the number of hidden residues determined from the contour length increment.

Fig. 2 shows a hypothetical protein containing two intricately interacting domains (red/blue rectangles) being stretched

2 The abbreviations used are: SMFS, single-molecule force spectroscopy; AFM, atomic force microscopy; WLC, worm-like chain; GdmCl, guanidinium chloride; RRL, rabbit reticulocyte lysate; pN, piconewtons; FE, force-extension curve.
between arbitrary protein handles (black ovals). In this schematic, we assume that the protein handles are more mechanically stable than the protein containing the interacting red/blue domains. The stretching of this entire protein construct typically follows a worm-like chain model of polymer elasticity (35). The initial contour length of the whole folded protein construct, \( L_1 \), has contributions from the protein handles (black ovals, each providing a contour length of \( x \)) and from the N-terminus to C-terminus distance in the two-domain protein (providing a contour length of \( l_{1} \)).

When domain D1 unfolds and stretching of the entire polyprotein continues, the next contour length, \( L_2 \), will contain three contributions: the unchanged length of the protein handles (again, each providing a contour length of \( x \)), the length of unfolded residues from D1 (which contribute a contour length, \( U_1 \), equal to the number of residues times the length per residue), and the N-terminus to C-terminus distance of domain D2 (which, after being reoriented, now contributes a contour length of \( l_{2} \)). Therefore, the difference between these two contour lengths, the contour length increment \( L_{c1} \), contains contributions from the initial length, \( l_{1} \), of the entire multidomain protein, the contour length \( U_1 \) from the unfolded polypeptides from the last domain (Fig. 2, domain D2) and contour length contributions from the initial length (Fig. 2, length \( l_{2} \)). That is, the final domain can be considered without any further contributions from reorientation because there are no further structural domains (e.g., \( I_3 / H_{11005} \) in Fig. 2). Therefore, the calculation for contour length increments in single-domain proteins is equivalent to the calculation for the final domain of a multidomain protein.

The accuracy of this model is determined by residuals. The residuals are the errors in this model associated with either U-errors, which are caused by incorrect assumptions about the boundaries of interacting domains, or I-errors, which are true deviations from the real native structure after a partial domain unfolds (i.e., it is possible that the unfolding of a domain of a multidomain protein causes a total reconfiguration to which there is no crystal structure standard to compare). The residual of a domain is given by the absolute difference between the measured contour length increment and the expected contour length increment, as we demonstrate below, limiting the precision of domain boundaries estimation.

In general, the contour length increment provided by unfolding of domain \( i \) in a multidomain protein (\( L_{ci} \)) is equal to the contour length from the unfolded residues in domain \( i \), plus the N-C length of the remaining protein after unfolding domain \( i \), minus the N-C length of the protein before unfolding of domain \( i \). The final domain has a final contour length increment that contains contributions from only the unfolded polypeptide from the last domain (Fig. 2, domain D2) and contour length contributions from the initial length (Fig. 2, length \( l_{2} \)).

The accuracy of this model is determined by residuals. The residuals are the errors in this model associated with either U-errors, which are caused by incorrect assumptions about the boundaries of interacting domains, or I-errors, which are true deviations from the real native structure after a partial domain unfolds (i.e., it is possible that the unfolding of a domain of a multidomain protein causes a total reconfiguration to which there is no crystal structure standard to compare). The residual of a domain is given by the absolute difference between the measured contour length increment and the expected contour length increment, as we demonstrate below, limiting the precision of domain boundaries estimation.

The contour length increments determined from the FE curve in AFM stretching measurements of a multidomain protein directly relate to the lengths of the polypeptide chains folded within the domains and also to their geometrical arrangement within the protein. A hypothetical protein containing two intricately interacting domains (red/blue rectangles) is being stretched between arbitrary protein handles (black ovals). The hypothetical force-extension curve is shown above each molecular schematic. The contour length increment upon each domain unfolding, \( L_{ci} \), is determined by the length of unfolded residues from each domain, \( U_i \), as well as by the gain in length because of the reorientation of the domains that remained folded within the multidomain protein.

**Figure 2.** The contour length increments determined from the FE curve in AFM stretching measurements of a multidomain protein directly relate to the lengths of the polypeptide chains folded within the domains and also to their geometrical arrangement within the protein. A hypothetical protein containing two intricately interacting domains (red/blue rectangles) is being stretched between arbitrary protein handles (black ovals). The hypothetical force-extension curve is shown above each molecular schematic. The contour length increment upon each domain unfolding, \( L_{ci} \), is determined by the length of unfolded residues from each domain, \( U_i \), as well as by the gain in length because of the reorientation of the domains that remained folded within the multidomain protein.
length increment, which is determined by the unfolding length of the corresponding domain and the initial length of the protein before and after unfolding of the domain (Fig. 2). The total residual is the sum of the residuals from all domains.

The domains that correspond to mechanical unfolding can then be determined by finding the correct domain partitioning that minimizes the total residual. This can be done by iteratively solving for the total residuals of all possible sets of domain boundaries and selecting the domain boundaries that provide the minimum total residual. The number of domains is always determined by the number of unfolding events, so the only parameters being determined are the residues that are at the interface of each domain.

Applying this model to determine the domain boundaries is important because gross errors will occur if the reorientation is not taken into account. Luciferase unfolding is a test case that exemplifies this scenario. Our force spectroscopy data indicate that Luciferase unfolds in three domains (no ligands), four domains (+ATP), or five domains (+ATP and +Luciferin). Fortunately, there is a model of the apo form of Luciferase (PDB code 1BA3) and a model of the ligand bound form (PDB code 2D1S) that can be used for calculating residuals from all possible domain boundaries. After determining boundaries that minimize the total residuals, we find that accounting for the reorientation gives an average total residual of 0.49 nm (apo, 0.35 nm; +ATP, 0.66 nm; +ATP +Luciferin, 0.47 nm). This can be roughly interpreted to mean that the model is able to account for all but approximately two residues. Conversely, by not accounting for reorientation (i.e. considering each interacting domain as a single domain), the average total residual is 6.3 nm (apo, 1.78 nm; +ATP, 6.96 nm; +ATP +Luciferin, 10.1 nm), which means that the model cannot account for ~17 residues, which is quite substantial considering that some domains are <100 residues. These errors increase when more domains are present because the reorientation would greatly affect each contour length increment. Therefore, we believe that, for multidomain proteins like Luciferase that have large interacting domains, it is absolutely crucial to account for reorientation when measuring mechanical domains by force spectroscopy to provide accurate estimates of the domain boundaries.

**Simulations**—Structure-based models were generated using the SMOG web server (36) using a Cα contact potential (37, 38). In this model, all residues are modeled as a single pseudoatom. The ATP ligand and Luciferin ligand were modeled as peptides that geometrically shared the same space as the ligands in the crystal structure to ensure that the same contacts were preserved when converted to a Cα model. The temperature was chosen as the folding temperature of the model Luciferase, T = 140 (Fig. 5). Simulations were conducted using GROMACS 4.5.5 (39) by pulling on the N terminus with reference to the C terminus at various speeds (0.1–1 nm/ns) and a spring constant of 6 pN/nm.

**RESULTS**

For mechanical unfolding and refolding experiments, we developed a Luciferase construct at the DNA level containing three I27 domains at the N terminus and four I27 domains at the C terminus (Fig. 3A) using the I27 domains only as a single-molecule fingerprint and handles for pickup (see “Experimental Procedures”). The expressed I27,-Luciferase-I27, protein construct still had native enzymatic activity (Fig. 1) and was used directly for AFM pulling measurements at constant velocity (see “Experimental Procedures”). The Luciferase protein unfolds in three distinct stages (Fig. 3B). The first peak (Peak 1, blue) has the smallest contour length increment ($\Delta L_c = 35.8 \pm 1.0$ nm, mean ± S.E., n = 548) and the lowest unfolding force ($F_u = 20.5 \pm 0.4$ pN, mean ± S.E., n = 548) (Fig. 3C). The second peak (Fig. 3C, Peak 2, green) increases in force and contour length increment ($\Delta L_c = 71.9 \pm 1.0$ nm, $F_u = 35.0 \pm 0.5$ pN, mean ± S.E., n = 548), and the third peak (Fig. 3C, Peak 3, red) has the highest force and the largest contour length increment ($\Delta L_c = 86.6 \pm 1.1$ nm, $F_u = 50.5 \pm 0.6$ pN, mean ± S.E., n = 548). These forces were determined at a constant loading rate of ~3900 pN/s. This was the only pathway detected for unfolding because the peak locations and relative unfolding forces followed the same pattern in all recordings, although, sometimes, the first peak in the force-extension trace was not detected or was possibly masked by breaking nonspecific interactions between the AFM tip and the substrate. To identify the
The origin of these peaks, we used a combination of computer models and experimental techniques.

The Mechanical Unfolding Pathway of Luciferase Proceeds from the C Terminus to the N Terminus—We first looked at the large amplitude fluctuations and dynamics of Luciferase using an elastic network model to probe the normal modes of motion (40). The lowest non-zero global modes from the normal mode analysis were calculated using the AD-ENM tool (41) using a model of the apo Luciferase. Models of Luciferase showed that the dominant mode of motion was the rotation and hinge motion of the C-terminal domain. There are 102 residues involved in this harmonic motion, which corresponds very well to the number of residues predicted from the contour length of the first peak: (32.5 nm / 3.3 nm initial length) / 0.37 nm/residue = 97 residues. Therefore, this model predicts that the first peak corresponds to the C-terminal domain. All dominant modes of motion corresponded to the C-terminal domain, which indicates that the large N-terminal domain of the protein (residues 1–450) does not have an obvious mechanical separation and cannot directly determine the origins of peak 2 or peak 3.

We also conducted steered molecular dynamic simulations on Ca models (42) of the apo form of the crystal structure of Luciferase at the folding temperature (Fig. 5 and “Experimental Procedures”). We used an open conformation structure to model the apo conformation (43) of Luciferase and the closed conformation (44) to model the ligand-bound unfolding pathways. Missing residues in the x-ray crystal structure were modeled using the SWISS-MODEL repository (45) using models 1ba3A (open) and 2d1sA (closed) (46). The temperature was chosen so that both the unfolded state and the folded state were equally populated at long time scales (~200 ns, 1 week of simulation time), as shown in Fig. 5. Steered molecular dynamic results showed excellent agreement with the AFM-FE curve of the apo form of Luciferase (Fig. 4, A and B, left panel) when simulated at the folding temperature (Fig. 5), suggesting that Luciferase stretching results in sequential unfolding of the C-terminal domain, the middle domain, and the N-terminal domain, corresponding to peaks 1, 2, and 3, respectively (Fig. 3B). Similar results were obtained for pulling on the C terminus with respect to the N terminus.

Because the Luciferase ligands (ATP and Luciferin) bind to residues within 250 amino acids and 450 amino acids, AFM and steered molecular dynamic experiments with ligands should only affect the unfolding of the middle of Luciferase. We performed pulling experiments of the Luciferase construct in the presence of its ligands, MgATP or MgATP with Luciferin (Fig. 3). In the presence of MgATP, there was a mixed population of AFM force-extension recordings that either corresponded to
Luciferase Refolding Rescued by Chaperones

FIGURE 6. Left panel, the mean ± S.E. of each unfolding force is shown for each condition and each peak. The unfolding forces of different peaks do not differ across different conditions. The unfolding force increases steadily from peak 1 to peak 2 and then decreases for peak 3. The presence of the ligands in the unfolding pathway only has the effect of splitting peak 2 and creating new peaks (peak 2’ or peak 2” and peak 2”) because the unfolding forces for peaks 1, 2, and 3 are not statistically different. Right panel, The mean ± S.E. of each contour length increment from each peak is shown for each condition and each peak.

the apo type (Fig. 4A, left panel) or a recording that had an additional peak, 2’, immediately after peak 2 (Fig. 4A, center panel). Interestingly, a third peak formed after peak 2’ appeared when Luciferin and MgATP were added, denoted peak 2” (Fig. 4A, right panel). For either condition, the only difference between any force-extension recordings of the unfolding of Luciferase is the added presence/absence of peak 2’ and/or peak 2” (Fig. 4C). The unfolding forces across different conditions for peak 1, peak 2, peak 2’, and peak 3 are comparable at similar loading rates because the null hypothesis that the force distributions have the same mean cannot be rejected after two-sample Student’s t test (Fig. 6). Because ligands bind to the middle of Luciferase and only affect the unfolding of the domain that corresponds to peak 2 in the force-extension trace, we hypothesize that peak 2 corresponds to the unfolding of the middle of Luciferase.

The observation that the unfolding force of peak 3 is much less than the unfolding force of the preceding peak 2” strongly suggests that unfolding occurs from one end to the other, as also suggested by our computer simulations. If the unfolding did not occur from end to the other, then that part of the protein would be unfolded in the center, which would mechanically separate two domains. If two domains are mechanically separated (by the unfolded residues), then the first to unfold would be the domain that has less mechanical stability. Because this is not the case and because the unfolding of peak 2” occurs at a much higher force than that of peak 3, then the unfolding must not separate the domains and must happen from one end to the other. We hypothesize that peak 2” occurs at a higher unfolding force than peak 3 because the domain corresponding to peak 3 is stabilized by the domain corresponding to the unfolding event from peak 2”. When the domain corresponding to peak 2” is unfolded, the domain corresponding to peak 3 loses some of its mechanical stability and unfolds at force lower than peak 2”.

We also did coarse-grained simulations of models of Luciferase with ligands that matched well the results of the experimental unfolding. The closed conformation of Luciferase (44) was used to model the ligand-bound unfolding pathways. The unfolding simulation of Luciferase with a modeled ATP ligand split peak 2 into a doublet of peaks 2 and 2’, and the simulations of models of Luciferase and a modeled ATP ligand and a modeled Luciferin ligand split peak 2 into a triplet of peaks 2, 2’, and 2” (Fig. 3). The forces do not compare well to the experiment because the unrealistic forces of the simulations may be due to the simplistic contact model inherent in Ca model potentials. However, the relative effects compare very well with the experiment and again indicate that the unfolding occurs in subdomains from the C terminus to the N terminus, which is consistent with the simulations of the apo form and with the results of the AFM measurements.

To further examine the unfolding pathway, we performed experiments with denaturant guanidinium chloride (GdmCl). Previous groups have measured equilibrium unfolding in denaturants using tryptophan fluorescence and found that specific domains unfold at various concentrations of GdmCl on the basis of the quenching of the tryptophan fluorescence (24, 26, 47). Notably, Wang et al. (47) made single-tryptophan mutants of Luciferase with tryptophan residues placed at various locations of Luciferase, which allowed them to probe equilibrium unfolding of domains independently. These studies found that many of the mutated tryptophan residues located in the middle domain of Luciferase are quenched at lower denaturants than the tryptophan residues located in the N-terminal domain (Fig. 6B), indicating that only the middle domain unfolds at low denaturant concentrations (24, 26, 47). It is difficult to determine the stability of the C-terminal domain because there was only one study that determined the effect of GdmCl on a tryptophan in the C-terminal domain.

We hypothesized that, because the N-terminal domain remains folded at low concentrations of denaturant and the middle domain unfolds at low concentrations of denaturant, AFM measurements of Luciferase in low concentrations of denaturant would still show peak 3 but not peak 2. When denaturant was added, we found that a proportion of force-extension curves contained peaks 1, 2, and 3, and another proportion only had peak 3 (Fig. 7B). The recordings that only had peak 3 had a long initial length, which would indicate that the domains corresponding to peak 1 and peak 2 are unfolded under these conditions. This result further confirms our hypothesis that the unfolding of the N-terminal domain gives rise to peak 3.

The proportion of FE recordings that had all peaks was concentration-dependent and decreased to about 50% at 250 mM GdmCl. This trend follows very closely the equilibrium unfolding of the middle domain of Luciferase measured by tryptophan fluorescence (Fig. 7C). We also noted that the presence of GdmCl did not seem to mechanically weaken the folded domains. This conclusion was made on the basis of the observation that the unfolding forces of peaks that were recorded in the presence of denaturant were indistinguishable from unfolding forces without denaturant (Fig. 7A). This suggests that the denaturant acts to completely unfold and not simply mechanically weaken the domain. Interestingly, this might indicate that the signal from tryptophan fluorescence equilibrium unfolding may integrate two populations of Luciferase molecules: one that is in the native state and one that has the second domain denatured.
To further validate the origin of peak 3 being the N-terminal domain, we directly probed the N-terminal domain by creating a truncated version of Luciferase, Luc203, containing residues 1–203 (Fig. 8). The contour length increment of the corresponding unfolding event in the force-extension trace is 49.4 ± 3.9 nm (mean ± S.D.). The force of Luc203 is very similar to the unfolding force of peak 3, which indicates a similar level of stability. Because it is unlikely that a truncation would result in a higher force than the original protein and that peak 3 had the highest force of any peak, it is very likely that the similar stability of Luc203 reflects that the N-terminal domain corresponds to peak 3. The difference in contour length increment (84.3 nm in the full unfolding pathway for peak 3 of Luciferase and 49.4 nm for the truncated Luc203) can be explained by the long unstructured coils left over from the truncation (Fig. 8C). The first 22 residues and residues 172–203 in Luciferase form long coils

**FIGURE 7.** Unfolding experiments in the presence of denaturant. A, the unfolding forces for peak 2 and peak 3 are plotted without denaturant (blue) and with denaturant and showing all peaks (gold) and with denaturant but only showing Peak 3 (magenta). The forces do not differ statistically, implying that the denaturant acts simply to unfold the protein rather than destabilize and lower the mechanical unfolding force. B, aggregates of FE curves for luciferase in the presence of GdmCl with all three peaks appearing (gold) and with only the third peak present (magenta). C, the red, blue, and green lines are measurements from references (24, 26, 47) shown here for comparison with our own data (black). These colored lines show data from chemical denaturation and subsequent evaluation of tryptophan apparent fluorescence (normalized to 1 at 0 m GdmCl and normalized to 0 at 5 m GdmCl). Colors indicate the location of residues in the protein (red, N-terminal residues; green, middle of protein; blue, C-terminal domain of protein). SMFS measurements (black) counted the fraction of SMFS curves with peak 2.

**FIGURE 8.** Luciferase was truncated to residues 1–203 to determine whether the unfolding force was similar to peak 3 (which should correspond to the unfolding of the N-terminal domain). The unfolding FE curve is shown in A, and the unfolding force is not statistically different from peak 3 in the apo form (B). However, the contour length increment is much smaller because of the fact that large loops are left without their normal contacts, making them unstructured and not contributing to the unfolding event in the truncation, as shown in C.
that would normally interact with the middle domain of Luciferase (residues 230–400). Therefore, it is expected that these residues will be unstructured in the truncated form. Thus, assuming that these coils only contribute to the initial length, the contour length increment should be equal to the length contributed by the number of residues in the rest of the structure minus the N-C length of that structured fragment (149 residues × 0.37 nm/residue − 2.8 nm = 52.3 nm). Therefore, because the truncated Luciferase does not contain the full structure, it is expected to give only ~52.3 nm of a contour length increment, which agrees well within the measured contour length increment of 49.4 ± 3.9 nm.

Structural Origins of Domains Involved in the Mechanical Unfolding Pathway—We conclude that this combination of experimental and simulated forced unfolding shows unequivocally that Luciferase undergoes sequential unfolding, domain by domain, from the C terminus to the N terminus. The precise structural domains can then be accurately determined by using AFM-measured contour length increments and measures of initial length from the crystal structure. By correlating the AFM contour length increments to the number of residues gained during extension, although considering all possible domain boundaries (see “Experimental Procedures”), we determined that the C-terminal domain (which gives rise to peak 1) consists of residues 443–550 and that the N-terminal domain (which gives rise to peak 3) consists of residues 1–203, with a total residual of 0.4 nm for the apo form (Fig. 9, left panel). These results are similar to algorithm predictions (N-terminal domain, start-266; C-terminal domain, 434-end) that consider multidomain proteins as collections of self-contained cooperative units (48). These divisions are also very similar to the divisions determined during the simulated unfolding (Fig. 9).

Similar methods can be used to determine the structural location of the domain unfolding events corresponding to peak 2’ and Peak 2” in the FE profile by assuming four domains in the case of +ATP and five domains in the case of +ATP and +Luciferin (see “Experimental Procedures”). All possible interfaces were considered, and the residuals were minimized with respect to the interfacial residues. The residuals were 0.7 nm for the +ATP experiments and 0.5 nm for the +ATP, +Luciferin experiments, which indicates that there is good agreement between the theoretically determined domains and the experimentally determined contour length increments from those domains. Because all other peaks remained at the same position (Fig. 4C) and the same force (Fig. 9), these additional unfolding events must have originated from the structural rearrangement caused by ligand binding. Indeed, the boundary between peak 1 and peak 2 and the boundaries around peak 3 are similar in each condition (Fig. 8).

When ligands are added, peak 2 has a contour length increment of ~8 nm (Fig. 5, right). Using domain boundaries determined by minimizing residuals, this contour length increment comes from the unfolding of residues 422–448. These residues have an initial length of ~1.5 nm (after accounting for the reorientation of the multidomain protein), so its theoretical contour length increment would be 8.5 nm (27 residues × 0.37 nm/residue − 1.5 nm), which corresponds very well with the expected contour length increment of ~8 nm. These residues correspond to a β-hairpin directly after the C-terminal hinge. The force required to rupture a single β-hairpin is ~25–50 pN, which is consistent with the force of peak 2 (~30 pN) (49). In the apo form, we expect that this is also the origin of peak 2 and that this unfolding event triggers a quick and subsequent unfolding of the rest of the middle of the protein (residues 248–422), which is otherwise stabilized by ligands in the ligand conditions. The structure that gives rise to peak 2’ in force-extension curves with ATP is likely the coil loop, which may make contacts with an unhydrolyzed ATP, and the stabilization of Luciferin in the middle of the protein likely gives rise to peak 2”.

This study is being followed up by mutagenesis to confirm these structural origins.

Separated Domains of Luciferase Refold Robustly—First we examined only the N-terminal domain using the truncated Luciferase containing only residues 1–203 (Luc203). The mechanical stability of the truncated Luciferase in stretching measurements was similar to the stability of N-terminal domain in the full-length protein, as determined by peak 3. Therefore, we assumed that the mechanical behavior of the isolated N-terminal domain is similar to its behavior in the complete protein. After a single molecule of Luc203 was extended, it was retracted and allowed to refold for 10 s, and then subsequent pulses of extension and retraction were performed. On subsequent extension pulses, a similar peak appeared, indicating that the truncated N-terminal domain of Luciferase refolds (Fig. 10, 2–5). The refolding of the truncated Luciferase occurred in 56% of subsequent extension pulses (n = 102, 12 molecules) with a time delay of 10 s.

Next we selectively unraveled Luciferase domains of the full Luciferase protein by controlling the extension of stretching and tested their ability to refold from the mechanically relaxed state. Using the entire Luciferase protein, we performed similar refolding experiments, first by selectively unfolding the C-terminal domain (Fig. 11A), which corresponds to peak 1 in the full unfolding of Luciferase. Correct refolding produces a similar peak upon the next pulse of unfolding (Fig. 11A, 2–5). When unfolding and refolding only the C-terminal domain or the middle domain (peaks 1 and 2, extending to 40 nm or 60–80 nm, respectively), it was observed that the peaks corresponding to unfolding of one of those domains reappeared in subsequent

FIGURE 9. The unfolding domains corresponding to the peaks in the FE curves were calculated (see “Experimental Procedures”) and are shown as Experiment (top panel). The domain boundaries from Simulation come from the unfolding trajectories in the steered molecular dynamic simulations. Bottom panel, The experimental boundaries are colored on the relevant structures (open for Apo and closed for the ligand bound).
pulling cycles, indicating refolding in about 78% of subsequent extension pulses ($n = 23$, six molecules). An estimate of the folding time can be obtained by decreasing the time delay between pulses. Refolding occurred when separating pulses by a time delay for 2 s, indicating a lower bound of the refolding rate of about $\sim 0.5 \text{ s}^{-1}$. We determined that, throughout these partial unfolding/refolding cyclic measurements, the N-terminal domain remained folded because the complete extension of the protein following these measurements captured its characteristic peak.

**Complete Unfolding of Luciferase Prevents Refolding unless Provided with Chaperones**—When Luciferase is fully stretched (thus completely unfolded) (extension 150–200 nm, showing peaks 1, 2, and 3) and then relaxed, the original unfolding peaks are never fully recovered in subsequent stretching cycles (Fig. 11B, I–5), suggesting that refolding is inhibited in these experiments. Instead, each pulse after the initial unfolding pulse shows irregular high-force events (100–200 pN). This result is observed even with a long time delay of 120 s ($n = 68$, 30 molecules). This indicates an upper bound for the refolding rate of $\sim 0.006 \text{ s}^{-1}$, but, because refolding was never fully observed, it is likely that this is a liberal upper bound.

We also performed refolding simulations analogous to the AFM experiments. The full apo model of Luciferase was extended only to unfold a single domain at a time (indicated by peak 1, peak 2, or peak 3), at which point the molecule was retracted, and then the force was taken off to allow complete relaxation and refolding. When the N-terminal domain was left intact (extending past either peak 1 or peak 2), Luciferase refolded in $< 3$ ns. When refolding from completely denatured Luciferase, the refolding occurs domain by domain from the N-terminal domain to the C-terminal (12 of 13 simulations). From the unfolded state, the mean time for the N-terminal domain to fold is $79 \pm 41$ ns, after which the middle domain folded in $3 \pm 2$ ns and then the C-terminal domain refolded in $9 \pm 6$ ns (mean $\pm$ S.D.). Therefore, in simulations, the refolding rate when the N-terminal domain is folded is over 20 times faster than when refolding from completely denatured protein. Both simulations and experiments show that Luciferase refolds robustly and quickly when the N-terminal domain is folded and

FIGURE 10. Refolding experiments of the truncated Luciferase construct, I27–Luc203–I27, (where Luc203 contains only residues 1–203), showing extension (red) and retraction (blue) for sequential measurements (black is a template molecule of the unfolding of truncated Luciferase). The time between extension pulses is 10 s, which is enough for I27 domains to refold 100% of the time and also allows Luc203 to refold most of the time (correct Luc203 refolding events are marked by asterisks).

FIGURE 11. Refolding experiments of the Luciferase construct showing extension (red) and retraction (blue) for sequential measurements (black is a template molecule). A, Luciferase is stretched only to the extension corresponding to the unfolding event of peak 1 or peak 2 and then relaxed and restretched with a time delay of 10–30 s, which showed the robust reappearance of peak 1 and peak 2. These peaks occur because of the unfolding of the C-terminal and middle domains, indicating that refolding occurs with a folded N-terminal domain. B, Luciferase is fully unfolded, and then cyclic unfolding/refolding measurements are performed after a delay of 30–180 s. In each cyclic measurement, the peaks do not reappear, and, instead, large nonspecific forces predominate, indicating stable misfolded states of Luciferase. C, Luciferase is fully unfolded in the presence of chaperones in the form of RRL, and subsequent pulses show peak reappearance, indicating successful partial refolding of Luciferase and no strong stable misfolds. Inset, the addition of RRL chaperones provides a statistically significant increase in the proportion of subsequent unfolding FE curves with refolded domains. Correct refolding events are marked by asterisks.
Luciferase Refolding Rescued by Chaperones

(A) First Extension

![Diagram](A)

(B) -RRL, subsequent extension

![Diagram](B)

(C) +RRL, subsequent extension

![Diagram](C)

FIGURE 12. The force peaks from refolding experiments are plotted by their unfolding forces and their contour length increments. In each plot, blue contour lines indicate where 90% of the native Luciferase peaks lie. Events that refold to the native structure should have subsequent peaks that correspond to these contours. The peaks of the first extension (no refolding) fall directly into these contours (A). Without chaperones (B), the peaks fall mostly outside and above the contours because they are mostly nonspecific and high-force events. With the addition of chaperones (C), the refolding peaks fall mostly inside contours, which indicates native-like refolding. The proportion of recordings with at least two refolding events were calculated for the full extension with and without chaperones and are plotted in Fig. 11C, inset.

that the refolding time is much longer when Luciferase is completely unfolded (molecular dynamics results).

To determine whether the fully unfolded Luciferase can be rescued by chaperones, we performed the AFM refolding experiments in the presence of rabbit reticulocyte lysate (RRL), which has been shown to renature Luciferase (28), because of the presence of chaperone proteins at high concentration. With the addition of RRL, Luciferase was fully extended and then relaxed for 120 s before extending again to monitor refolding in the second pulse. The FE curve of the subsequent pulses frequently showed similar peaks as the original, indicating that refolding was successful (Figs. 11C and 12). This effect was statistically significant because the number of force-extension curves with at least two refolded domains increases from 13% without RRL (95% confidence interval, 6–24%, n = 68 on 30 molecules) to 46% with RRL (95% confidence interval, 33–61%, n = 54 on 16 molecules). This result suggests that the high-force peaks without RRL are due to misfolded states stabilized by non-native contacts and that chaperones inhibit these contacts helping the unfolded and relaxed protein to recover its native structure. We are actively investigating the specific chaperones that enable refolding.

DISCUSSION

Firefly luciferase (Luciferase) is homologous to AMP-forming ligases and nonribosomal peptide synthetases and it is one of the first described large, multidomain proteins shown to undergo cotranslational folding and chaperone-assisted refolding. Important questions remain about how cotranslational folding prevents misfolding, how chaperones effectively refold the protein, and why Luciferase forms such a stable misfolded state. To address these questions, we used SMFS and computer simulations to selectively denature the C-terminal domain, the middle domain, or the full protein, including the N-terminal domain (corresponding to peak 1, peak 2, and peak 3 in the unfolding force-extension curve), and performed refolding studies on individual protein molecules.

The refolding studies on individual molecules with AFM has the drawback of possibly introducing unanticipated surface protein effects because the molecule of interest is attached to the surface nonspecifically. We believe our results circumvent this issue for three main reasons. All refolding measurements are done after retracting the tip away from the surface by at least 15 nm before performing any refolding; all constructs are flanked by I27 domains that should provide an additional buffer of 12–16 nm if tethered at the ends; and there are drastic differences between the refolding in different experiments (truncated Luciferase, fully denatured native Luciferase with/without chaperones, and partially denatured Luciferase), which would likely not be the case if nonspecific surface effects were prevalent. However, we are further investigating possible effects using a variety of surfaces.

Our first major observation is that Luciferase is able to prevent misfolding and refold robustly when the N-terminal
domain is left folded. The C-terminal and middle domains of Luciferase can undergo many cycles of unfolding and refolding without forming observable non-native intermediates as long as the N terminus is folded. Their refolding rate (lower limit \(\sim 0.5 \text{ s}^{-1}\)) is also much faster than the typical cotranslational folding rate of Luciferase (\(\sim 0.005 \text{ s}^{-1}\)). This suggests that the bottleneck of Luciferase folding is the folding of its N-terminal domain.

To probe the folding of the N-terminal domain by itself, we used a truncated domain consisting of only the first 203 residues (Luc203). This domain refolded robustly but not as efficiently as the C-terminal domains in the presence of the folded N-terminal domain. This experiment was performed on a truncated protein that may have had different properties than the native protein, and we are further investigating this by performing refolding with a pulling geometry that directly affects the N-terminal domain within the native protein.

The refolding of the full protein was prevented when Luciferase was fully denatured mechanically, which has been observed previously using chemical denaturants (24, 25). However, instead of attributing the lack of refolding to aggregation, we suggest that it is due to misfolding. Because we observed that refolding can occur in an isolated N-terminal domain (by protein truncation) and that refolding can occur in the C-terminal domains if the N-terminal is folded, we suggest that the misfolding is due to the interaction between the unfolded N-terminal and unfolded C-terminal residues of a single protein. Such misfolding is easily prevented in the cell, where the cotranslational folding of the N-terminal domain prevents interaction with the C-terminal residues, thereby preventing putative very stable, non-functional intermediates (50).

Finally, we observed that adding chaperones in the form of cell lysate can significantly increase the refolding of fully denatured Luciferase, as reported previously in the literature, using bulk methods (28). However, our results with partial denaturation and the truncated Luciferase also suggest a specific mechanism by which chaperones could act on Luciferase. Because partial unfolding does not require chaperones for efficient refolding, chaperones must act to specifically enable folding of the N-terminal domain, which should then catalyze the refolding of the rest of the protein. Whether this action is achieved by association of chaperones with the N-terminal domain and separating it from C-terminal residues or sequestering C-terminal residues from the N-terminal residues remains to be examined. The presence and action of these chaperones seem to emulate the conditions provided by the ribosomal exit tunnel, assisting the nascent chain in the vectorial nature of cotranslational folding. In this way, both the ribosome and the chaperone would allow the sequential folding to avoid the kinetic traps inherent in the energy landscape of Luciferase. It is likely that folding of other multidomain proteins may exploit a similar mechanism.

REFERENCES

1. Dobson, C. M. (2003) Protein folding and misfolding. Nature 426, 884–890
2. Fersht, A. (1999) Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding, p. 1, Macmillan, London
3. Bartlett, A. I., and Radford, S. E. (2009) An expanding arsenal of experimental procedures yields an explosion of insights into protein folding mechanisms. Nat. Struct. Mol. Biol. 16, 582–588
4. Englander, S. W., Mayne, L., and Krishna, M. M. (2007) Protein folding and misfolding: mechanism and principles. Q. Rev. Biophys. 40, 287–326
5. Onuchic, J. N., and Wolynes, P. G. (2004) Theory of protein folding. Curr. Opin. Struct. Biol. 14, 70–75
6. Eaton, W. A., Muñoz, V., Hagen, S. J., Jas, G. S., Lapidus, L. J., Henry, E. R., and Hofrichter, J. (2000) Fast kinetics and mechanisms in protein folding. Annu. Rev. Biophys. Biomol. Struct. 29, 327–359
7. Schuler-Furman, O., Wang, C., Bradley, P., Misura, K., and Baker, D. (2005) Progress in modeling of protein structures and interactions. Science 310, 638–642
8. Dill, K. A., and MacCallum, J. L. (2012) The protein-folding problem, 50 years on. Science 338, 1042–1046
9. Onuchic, J. N., Luthey-Schulten, Z., and Wolynes, P. G. (1997) Theory of protein folding: the energy landscape perspective. Ann. Rev. Phys. Chem. 48, 545–560
10. Wolynes, P. (2005) Recent successes of the energy landscape theory of protein folding and function. Q. Rev. Biophys. 38, 405–410
11. 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
12. Peng, Q., and Li, H. (2008) Atomic force microscopy reveals parallel mechanical unfolding pathways of T4 lysozyme: evidence for a kinetic partitioning mechanism. Proc. Natl. Acad. Sci. U.S.A. 105, 1885–1890
13. Kaiser, C. M., Goldman, D. H., Chodera, J. D., Tinoco, I., Jr., and Bustamante, C. (2011) The ribosome modulates nascent protein folding. Science 334, 1723–1727
14. Kimchi-Sarfaty, C., Oh, J. M., Kim, I.-W., Sauna, Z. E., Calagno, A. M., Ambudkar, S. V., and Gottesman, M. M. (2007) A “silent” polymorphism in the MDR1 gene changes substrate specificity. Science 315, 525–528
15. Pinchi, M., Ziv, G., Riven, I., Cohen, S. S., Zohar, N., Barak, Y., and Haran, G. (2011) Single-molecule fluorescence spectroscopy maps the folding landscape of a large protein. Nat. Commun. 2, 493
16. Chakraborty, K., Chatila, M., Sinha, J., Shi, Q., Poschner, B. C., Sikor, M., Jiang, G., Lamb, D. C., Hartl, F. U., and Hayer-Hartl, M. (2010) Chaperonin-catalyzed rescue of kinetically trapped states in protein folding. Cell 142, 112–122
17. Zhou, H.-X., and Dill, K. A. (2001) Stabilization of proteins in confined spaces. Biochemistry 40, 11289–11293
18. Denburg, J. L., Lee, R. T., and McElroy, W. D. (1969) Substrate-binding properties of firefly luciferase: I. luciferin-binding site. Arch. Biochem. Biophys. 134, 381–394
19. Lee, R. T., Denburg, J. L., and McElroy, W. D. (1970) Substrate-binding properties of firefly luciferase: II. ATP-binding site. Arch. Biochem. Biophys. 141, 38–52
20. DeLuca, M., and McElroy, W. D. (1974) Kinetics of the firefly luciferase catalyzed reactions. Biochemistry 13, 921–925
21. Chang, K. H., Xiang, H., and Dunaway-Mariano, D. (1997) Acyl-adenylate motif of the acyl-adenylate/thioester-forming enzyme superfamily: a site-directed mutagenesis study with the Pseudomonas sp. strain CBS3 4-chlorobenzoate-coenzyme A ligase. Biochemistry 36, 15650–15659
22. Kleinkauf, H., and Von Döhren, H. (1996) A nonribosomal system of peptide biosynthesis. Eur. J. Biochem. 236, 335–351
23. Ozawa, T., Kihara, A., Sato, M., Tachihara, K., and Umezawa, Y. (2001) Split luciferase as an optical probe for detecting protein-protein interactions in mammalian cells based on protein splicing. Anal. Chem. 73, 2516–2521
24. Herbst, R., Gast, K., and Seckler, R. (1998) Folding of firefly (Photinus pyralis) luciferase: aggregation and reactivation of unfolding intermediates. Biochemistry 37, 6586–6597
25. Herbst, R., Schäfer, U., and Seckler, R. (1997) Equilibrium intermediates in the reversible unfolding of firefly (Photinus pyralis) luciferase. J. Biol. Chem. 272, 7099–7105
26. Frydman, J., Erdjument-Bromage, H., Tempst, P., and Hartl, F. U. (1999) Co-translational domain folding as the structural basis for the rapid de novo folding of firefly luciferase. Nat. Struct. Biol. 6, 697–705
27. Svetlov, M. S., Kommers, A., Kolb, V. A., and Spiri, A. S. (2006) Effective cotranslational folding of firefly luciferase without chaperones of the
Hsp70 family. Protein Sci. 15, 242–247
28. Schumacher, R. J., Hurst, R., Sullivan, W. P., McMahon, N. J., Toft, D. O., and Matts, R. L. (1994) ATP-dependent chaperoning activity of reticulocyte lysate. J. Biol. Chem. 269, 9493–9499
29. Bertz, M., and Rief, M. (2008) Mechanical unfoldons as building blocks of maltose-binding protein. J. Mol. Biol. 378, 447–458
30. 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
31. Shank, E. A., Cecconi, C., Dill, J. W., Marqusee, S., and Bustamante, C. (2010) The folding cooperativity of a protein is controlled by its chain topology. Nature 465, 637–640
32. Stigler, J., Ziegler, F., Gieseke, A., Gebhardt, J. C., and Rief, M. (2011) The complex folding network of single calmodulin molecules. Science 334, 512–516
33. Kim, M., Abdi, K., Lee, G., Rabbi, M., Lee, W., Yang, M., Schofield, C. J., Bennett, V., and Marszalek, P. E. (2010) Fast and forceful refolding of stretched α-helical solenoid proteins. Biophys. J. 96, 3086–3092
34. Dietz, H., and Rief, M. (2006) Protein structure by mechanical triangulation. Proc. Natl. Acad. Sci. U.S.A. 103, 1244–1247
35. Bustamante, C., Marko, J. F., Siggia, E. D., and Smith, S. (1994) Entropic elasticity of phage DNA. Science 265, 1599–1600
36. 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
37. Lammert, H., Schug, A., and Onuchic, J. N. (2009) Robustness and generalization of structure-based models for protein folding and function. Proteins 77, 881–891
38. 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
39. Pronk, S., Pääl, 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
40. Zheng, W., Brooks, B. R., and Thirumalai, D. (2006) Low-frequency normal modes that describe allostERIC transitions in biological nanomachines are robust to sequence variations. Proc. Natl. Acad. Sci. U.S.A. 103, 7664–7669
41. Zheng, W., and Doniach, S. (2003) A comparative study of motor-protein motions by using a simple elastic-network model. Proc. Natl. Acad. Sci. U.S.A. 100, 13253–13258
42. Clementi, C. (2008) Coarse-grained models of protein folding: toy models or predictive tools? Curr. Opin. Struct. Biol. 18, 10–15
43. Franks, N. P., Jenkins, A., Conti, E., Lieb, W. R., and Brick, P. (1998) Structural basis for the inhibition of firefly luciferase by a general anesthetic. Biophys. J. 75, 2205–2211
44. Nakatsu, T., Ichiyama, S., Hiratake, J., Saldanha, A., Kobashi, N., Sakata, K., and Kato, H. (2006) Structural basis for the spectral difference in luciferase bioluminescence. Nature 440, 372–376
45. Guex, N., Peitsch, M. C., and Schwede, T. (2009) Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: a historical perspective. Electrophoresis 30, S162–S173
46. Kiefer, F., Arnold, K., Künzli, M., Bordoli, L., and Schwede, T. (2009) The SWISS-MODEL Repository and associated resources. Nucleic Acids Res. 37, D387–D392
47. Wang, W.-Q., Xu, Q., Shan, Y.-F., and Xu, G.-J. (2001) Probing local conformational changes during equilibrium unfolding of firefly luciferase: fluorescence and circular dichroism studies of single tryptophan mutants. Biochem. Biophys. Res. Commun. 282, 28–33
48. Porter, L. L., and Rose, G. D. (2012) A thermodynamic definition of protein domains. Proc. Natl. Acad. Sci. U.S.A. 109, 9420–9425
49. Sapra, K. T., Damaghi, M., Köster, S., Yildiz, O., Kühlbrandt, W., and Muller, D. J. (2009) One β hairpin after the other: exploring mechanical unfolding pathways of the transmembrane β-barrel protein OmpG. Angew. Chem. Int. Ed. Engl. 48, 8306–8308
50. Clark, P. L. (2004) Protein folding in the cell: reshaping the folding funnel. Trends Biochem. Sci. 29, 527–534