Published in final edited form as:

J Mol Biol. 2009 September 11; 392(1): 218–227. doi:10.1016/j.jmb.2009.06.072.

Chromophore packing leads to hysteresis in GFP

Benjamin T. Andrews, Melinda Roy, and Patricia A. Jennings*

Department of Chemistry & Biochemistry, University of California at San Diego, 9500 Gilman Dr., La Jolla, CA, 92093-0375

Abstract

Green fluorescent protein (GFP) possesses a unique folding landscape with a dual basin, leading to the hysteretic folding behavior observed in experiment. While theoretical data do not have the resolution necessary to observe details of the chromophore during refolding, experimental results point to the chromophore as the cause of the observed hysteresis. Using NMR spectroscopy, which probes at the level of the individual residue, the hysteretic intermediate state is further characterized in the context of the loosely-folded native-like state $\{N_{iso}\}$ predicted in simulation. In the present study, several residues located in the lid of GFP indicate heterogeneity of the native states. Some of these residues show chemical shifts when the native-like intermediate $\{N_{iso}\}$ responsible for GFP's hysteretic folding behavior is trapped. Observed changes in the chromophore are consistent with increased flexibility or isomerization in $\{N_{iso}\}$ as predicted in recent theoretical work. Here we observe multiple chromophore environments within the native state are averaged in the trapped intermediate, linking chromophore flexibility to mispacking in the trapped intermediate. The present work is experimental evidence for the proposed final “locking” mechanism in GFP folding forming an incorrectly or loosely packed barrel under intermediate (hysteretic) folding conditions.

Keywords

protein folding; energy landscape; proline isomerization; folding intermediate; fluorescent protein

Introduction

Green fluorescent protein (GFP) is a commonly used reporter both *in vivo* and *in vitro* due to its ability to fold and autocatalytically form a visually fluorescent chromophore.¹ Chromophore formation is linked to folding,² and is required for reporter applications. However, the time-scale of GFP folding limits its usefulness for following time-related phenomena. Thus, improving the rate and efficiency of GFP protein folding could improve its use as a reporter.³ Recent work on the folding of GFP details large barriers and hysteresis in folding⁴ linked to the chromophore⁵(Figure 1). Proline isomerization in the lid is linked to chromophore formation.⁵ The question remains, how does a formed chromophore lead to the hysteresis observed in folding?

The structure of sfGFP is a highly regular 11-stranded β -barrel with a central, kinked α -helix running through the center (Figure 2A & 2B).⁶ The chromophore forms from an autocatalytic reaction of the backbone involving cyclization, oxidation, and dehydration

*Corresponding Author: Patricia A. Jennings, pajennin@ucsd.edu, (858) 534-6417.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

reactions.^{7, 8, 9, 10, 11} Chromophore formation follows the kinking of the helix; hence, the chromophore does not cause the kinked helix.² Chromophore fluorescence depends on segregation from bulk solvent as provided by the barrel,¹² and requires rigidity in chromophore structure.¹³ However, despite being surrounded by an 11-stranded β -barrel, the chromophore is not always sufficiently rigid for fluorescence and may isomerize in a hula-twist (HT) motion (Figure 2C).

Chromophore isomerization in GFP was originally linked to the “blinking” observed in single molecule fluorescent protein (FP) studies.¹⁴ This is similar to the isomerization observed via crystal structures in photoactive yellow protein, another FP structurally distant from GFP,^{15, 16} and to the photoisomerization observed in DsRed, a red fluorescent protein structurally analogous to GFP.¹⁷ Fluorescence quenching in yellow protein and DsRed are the result of interconversion during a non-adiabatic crossing.^{14, 18} Several molecular dynamics (MD) studies have been performed to predict the details of isomerization within the β -barrel of GFP. These simulation studies consistently observe that the *cis* form found in the crystal structure is a lower energy form, as the *trans* form appears to lose some of the H-bonds situating it within the barrel.^{19, 20, 21} Model compound studies also observe the *cis* form to be 2.5 kcal/mol more stable than the *trans* form, with a 13 kcal/mol barrier in the neutral form.^{19, 21} However, while both *cis* and *trans* chromophore states are fluorescent, interconversion between the two is proposed to quench fluorescence.

Chromophore isomerization has not been experimentally observed within the barrel of GFP, although changes in the chromophore flexibility have been linked with fluorescent and non-fluorescent states.^{22, 23, 24} Work on sfGFP⁶ proposed chromophore isomerization as a potential source of hysteresis in GFP,⁵ and recent theoretical work supported that the final β -strand inserting into the barrel is the final step in folding.²⁵ The theoretically predicted loosely packed, native-like intermediate is referred to as $\{N_{iso}\}$, an intermediate dominated by an ensemble of structures likely with non-native proline and chromophore isomerizations. Taken together, these results are consistent with the final strand insertion trapping a poorly packed barrel and chromophore mis-isomerization. However, until now, there has been no probe to unambiguously follow the chromophore structure or isomerization within the barrel, and little experimental evidence to support this conjecture.

Probes of the GFP chromophore exist, but typically consist of chromophore fluorescence or absorbance, neither of which gives clear structural detail of the chromophore. While fluorescence may change based on the structure of the chromophore,²¹ other factors, such as the surrounding environment may also contribute changes and one may question whether the chromophore, the environment, or a likely combination of both may be causing a signal change. While NMR is a useful probe to assess the behavior of specific amino acid residues, ¹H spectra have been difficult to interpret in full length GFP due to insufficient resolution of ¹H protein signals.⁵ Work on model chromophores show changes in NMR based on the isomerization.²¹ This, in combination with peptide work has shown that homonuclear 2D NMR may be able to tease out structural changes in the protons of the chromophore.²⁶ Also, while not discussed here, Raman spectroscopy has potential in measuring the isomerization state of the chromophore within GFP.^{27, 28}

The current work uses NMR spectroscopy to compare the native and trapped intermediate states of GFP which are captured within the hysteresis zone (Figure 1). The trapped state is consistent with a kinetically stable, but poorly packed structure containing loose turns and a loose chromophore, similar to $\{N_{iso}\}$ as predicted in theoretical simulation. Residues which show shifts in a ¹H-¹⁵N heteronuclear single quantum coherence spectra (HSQC) spectrum²⁵ now also depict protein heterogeneity in the ¹H, ¹⁵N, and ¹³C chemical shifts of the peptide backbone of the native structure observed in triple resonance NMR

spectroscopy. Heterogeneity in additional residues is also pinpointed by triple resonance NMR. These residues reside in the lid of the barrel of GFP. Multiple structural states in the lid are further evidence of a loose structural state creating a kinetically trapped intermediate, resulting in the hysteresis observed in experimental data⁵ which had been previously predicted in simulation data.²⁵ Some of these residues have been linked to chromophore formation. Furthermore, we present the first evidence of changes in the chromophore structure within the hysteretic intermediate, which are linked to increased flexibility and rapid interconversion of the chromophore as we previously predicted. Our results are consistent with a native structure containing a twisted chromophore with two states, while the $\{N_{iso}\}$ structure is averaged between the two. It is harder to pack an already formed chromophore inside the barrel. Thus, we present experimental evidence that refolding with a pre-formed chromophore leads to a loose, incorrectly packed intermediate that is stable enough to observe hysteresis. This fits in with our predictions based on theoretical evidence.

Results & discussion

Structural Heterogeneity within the Native Ensemble

The GFP folding landscape displays long-lived intermediates,⁴ as well as hysteresis in equilibrium folding corresponding to structural changes detected by 1D ^1H NMR spectroscopic studies.⁵ In fact, more recent work using 2D ^1H - ^{15}N NMR spectroscopy observed structural changes linked to the prolines linked to chromophore formation.²⁵ In the current study, triple-resonance 3D NMR spectroscopy is used to observe additional structural perturbations in GFP. The ^1H , ^{15}N , ^{13}C chemical shifts are analyzed in order to assess structural perturbations of specific residues within the polypeptide backbone (Figure 3). Several residues in the native structure display conformational heterogeneity in Figure 3. The residues which display heterogeneity are mapped onto the crystal structure in Figure 4 are on the lid of the barrel.

These results show conformational heterogeneity in the “top” lid of GFP, where both the N and C termini are located. The structural changes previously observed in the lid of the trapped sample are consistent with these results²⁵ and map near prolines linked to chromophore formation.⁵ In addition, we now observe heterogeneity in the backbone of W57, a residue thought to be important in the optimal formation of the native chromophore. Unfortunately, the $\{N_{iso}\}$ trapped species was not amenable to 3D NMR experiments owing to poor solubility and weak NMR signal.

As a way to overcome NMR signal issues in the trapped species, we used hydrogen deuterium exchange (HDX) to determine if structural changes in the $\{N_{iso}\}$ result in differences in protection of amide exchange from secondary structure alteration (Figure 4). Generally, in both the native and trapped structures, residues within the β -barrel exchange slowly, i.e. have high protection factors while residues on the loops exchange very fast as a result of the lack of secondary structure and hydrophobic interaction (Figure 4). The overall similarity in H/D exchange is the result of a similar global structure. However, there are residue-specific differences in solvent exchange behavior between the native and trapped structure and the differences are quite telling. First, the trapped structure has several extra residues in the N-terminal helix which the observed rates of exchange are so fast that signals are gone by the first couple of time points. The same residues in the native state show exchange at least 10-fold slower than an intrinsic rate (protection factor (PF) <10). The vastly increased exchange rate for these residues in the trapped structure is a result of poor packing in the lid of the barrel. This is consistent with the loose structure that we previously predicted in simulation studies.²⁵ Several additional residues in the trapped structure that also display increased exchange with solvent are shown in Figure 4. The residues with

changes in exchange rate (>3) propagate through the lid, which are the result of a loose structure in the lid which can more readily exchange with the bulk solvent.

A Structural Look at the Chromophore and Implications for Isomerization & Interconversion

The GFP chromophore is made from a post-folding covalent modification which includes the loss of an amide proton, and thus requires analysis of $^1\text{H}^\alpha$ or $^1\text{H}^\beta$ or $^{13}\text{C}^{\alpha/\beta}$ resonances, for example, as there is no $^1\text{H}^N$ signal. Here, we present the aromatic region of homonuclear 2-D TOCSY spectra where the spin system of the chromophore in GFP is observed (Figures 5 & 6). Resonances from the chromophore spin system are now assigned and compare well with published assignments of peptide fragments²⁶ as well as *cis-trans* isomerization work on 1D ^1H spectra of model compounds.²¹

Comparison of the aromatic region of the TOCSY spectra of both the native chromophore and trapped $\{\text{N}_{\text{iso}}\}$ show that a majority of the peaks overlay (Figure 5). This is consistent with the experimental data from circular dichroism, fluorescence, and NMR, as well as simulation data. All indicate that the trapped species has a similar global structure to the native state.^{5; 25} However, the NMR spectra exhibit local, residue-specific changes which are directly attributed to the chromophore. The $\text{H}^\delta\text{H}^\epsilon$ interaction characteristic of the chromophore spin system does not change between the native and trapped states, consistent with an intact structure of the chromophore (Figure 5), and with the characteristic coupling on the phenyl ring. However, changes are observed in the H^β resonance (Figure 6), which lies on the link between the rings of the chromophore. In the native state, the $\text{H}^\epsilon\text{H}^\beta$ resonance shows two peaks, evidence of two rigid or slowly exchanging structures. However, in the trapped state, the $\text{H}^\epsilon\text{H}^\beta$ resonances are averaged, indicative of the chemical environment quickly changing between two states, or fast exchange. The exchange is likely between the two native states observed in the native structure to either a planar (non-fluorescent) state, or a flexible state. These results are consistent either with chromophore flexibility, which has been observed in Dronpa, or with chromophore isomerization,²³ which has been observed in model compound studies experimentally,^{19; 21} and in modeling studies within the β -barrel.^{18; 20; 29; 30}

Interplay between Lid and Internal Helix is Linked to the Chromophore

Superfolder GFP exhibits hysteresis (Figure 1) which is attributed to a structurally similar, but subtly different intermediate structure which contains a chromophore.⁵ Simulation data of the folding mechanism predicted that the final, rate limiting step in folding is packing the last strand into the β -barrel of GFP, packing the chromophore.^{25; 31} We proposed that this final step in folding was locking the loose barrel around the central α -helix creating a highly stable, locked structure once all strands are correctly placed. This requires precise packing of the interior residues which led to the experimentally observed extremely slow rate in the folding and the experimentally observed hysteresis.²⁵ Residues G35 and E90 were previously known to have different ^1H - ^{15}N HSQC chemical shifts in the trapped intermediate.²⁵ Now, changes in H/D exchange rates and HNCA show conformational heterogeneity extends further throughout the lid (Figure 4). In addition, P58 in sfGFP has been proposed to be part of the structural requirements for chromophore formation; double mutants involving P75 or P89 with P58 do not express or fold well. W57 exhibits heterogeneity associated with proline *cis-trans* isomerization in cyan FP.³² Here, we see evidence that P58 may be part of the proposed “triumvirate” of proline locks involved with chromophore formation. Again, top lid prolines (P75 & P89) appear to be conserved across 11-stranded β -barrel proteins that fluoresce, but are not present in the non-fluorescent, structurally analogous nidogen-1.³³ Furthermore, a recent split GFP reporter is able to form

a fluorescent chromophore without all the barrel strands. However, the portion of split GFP containing the chromophore also contains these prolines (P58, P75, and P89).³⁴

The loose GFP structure predicted from simulation data $\{N_{iso}\}$ before final structural locking has now been demonstrated by spectroscopic data. Theoretical data predict that under highly unfolding conditions, GFP could take either a direct route from native to unfolded, or populate a short-lived intermediate. Under intermediate unfolding conditions within the hysteresis zone, simulation data predicts that GFP must unfold via a long-lived intermediate.^{25, 35} Again, our experimental data here, showing a native-like, loose structure populated at intermediate unfolding (hysteretic) conditions, confirms the accuracy of our prediction. In addition, when GFP unfolding is followed by multiple probes, there is native state disruption at $[Gdn-HCl] < 1M$ following tryptophan fluorescence whereas chromophore fluorescence exists at $[Gdn-HCl] < 4M$, i.e. at more denaturing conditions.³⁶ Hence, the chromophore remains “locked down” and sequestered from the bulk solvent long after stable tertiary structure has been disrupted. Also, in the presence of trigger factor protein, with proline isomerase activity, the recovery of W57 fluorescence is accelerated,³⁶ linking these results to P58 isomerization, as we predicted.⁵ Finally, Enoki *et. al.*³⁷ reported initial packing of the tryptophan probe, followed by disruption of the packing, and then subsequently final packing to the native state. This again is consistent with a breathing motion which could be present in the final locking mechanism. Because of the unique structure of GFP having the chromophore-containing α -helix in the center of this locking barrel, there appears to be interesting consequences from this mechanism.

Chromophore Packing during Folding Leads to a Rough Energy Landscape

Once the chromophore is formed, the folding landscape of sfGFP becomes very rough⁵. The packing of the final strand into the barrel becomes problematic as the already formed chromophore creates structural restrictions during barrel packing due to the lowered flexibility of the chromophore side chains. A related FP, Dronpa, has conformationally mobile β -strands related to chromophore structure.²³ These results are similar to the flexibility found in GFP strands 3,7,8 and 10 as observed in NMR solvent exchange studies,^{4, 38} and are attributed to folding intermediates. These strands in GFP lie just outside the hydroxyl of the chromophore, and may sample different conformations in order to accommodate chromophore flexibility. We also observe resonances from the β -strands 3, 7, 8, and 10 in sfGFP with very broad line widths in HSQC spectra.²⁵ Broadened line widths relative to the size of the protein indicate an intermediate rate of exchange between different populations; this is consistent with populations shifting between different conformations. Furthermore, the interior of the barrel is not compatible with a planar chromophore.²⁹ Packing the barrel involves forcing a slight twist in the chromophore, which would otherwise be completely planar. Model chromophore compounds are able to isomerize, so chromophore isomerization likely occurs in the unconstrained, unfolded state (albeit slowly).²¹ Under proper conditions (e.g. the hysteretic refolding transition), the barrel may form around the inappropriate isomer form of the chromophore, which leads to a poorly packed barrel, resulting in the observed trapped intermediate.

In Figure 6, homonuclear 1H - 1H TOCSY NMR spectra show evidence of structural changes in the H_β of sfGFP's chromophore, consistent with conformational heterogeneity between the rings. The H_β in the chromophore lies on the carbon which undergoes a massive rearrangement during an HT isomerization of the chromophore. Clearly, there are changes in the chromophore between the trapped intermediate and native state, and chromophore packing has changed within the hysteretic folding zone (Figure 1). Theoretical data indicate that an HT chromophore isomerization is volume conserving, making it a possible pathway of isomerization or chromophore flexibility.¹⁸ Thus, within the trapped intermediate, we see chromophore flexibility (Figure 6) linked to the loose lid of the barrel (Figure 4 and ²⁵), the

link between the previously observed chromophore-dependent hysteresis⁵, and the theoretically predicted loose-state populated late in folding.²⁵

Final Locking Mechanism Problematic with Formed Chromophore

While landscape theory proposes that minimal frustration during folding is one “goal” during evolution, the protein must also be functional. Simulated and experimental data show that the functional loops or moieties lead to topological frustration in the folding landscape.^{39, 40, 41} GFP folding experiments in the absence of a formed chromophore have a smoother folding landscape, and *in vitro* folding in the presence of an already formed chromophore has a more rugged folding landscape. For example, in the absence of a formed chromophore, the tyrosine ring (Y66) within the chromophore would have much higher configurational entropy, and would more easily pack within the barrel. The physiological function of GFP in *Aequoria Victoria* is unknown and there is no evidence that *in vivo*, once folded and mature, GFP unfolds or undergoes any sort of major configurational changes. However, the highly stable barrel appears to segregate the chromophore from bulk solvent,¹² and maintains chromophore geometry to allow fluorescence.¹⁸ Therefore, the GFP chromophore may be a functional moiety which creates a limit on folding rate or robustness.

During *in vitro* refolding, a threshold of stability is required to correctly lock down the barrel around the formed chromophore, leading to the hysteresis observed at intermediate Gdn-HCl concentrations. When the threshold of stability is not met, GFP may sample an improperly packed barrel, with a mis-isomerized or flexible chromophore, leading to lower observed fluorescence and mispacking in the lid of the protein (Figure 4). It may be possible to follow changes within the barrel during folding through hydration NMR^{42, 43} to determine the location of structural waters. Raman spectroscopy, which can directly probe the isomerization state of the chromophore,^{27, 28} may potentially determine the differences between native and intermediate states.

In vitro refolding of GFP encounters a uniquely rugged landscape as the result of the barrel needing to fold around an already formed chromophore. In essence, there is a local minimum in folding landscape in which an incorrectly isomerized chromophore can be kinetically trapped, leading to the observed hysteretic behavior. In addition, heterogeneity in the lid results from a loose/conformationally flexible structure in residues within the hysteretic zone. As predicted earlier,⁵ the chromophore is heterogeneous but rigid within the barrel of folded GFP, and is loose in the weakly fluorescent, hysteretic intermediate.

Methods & materials

sfGFP Expression & Purification

Recombinant superfolder GFP (sfGFP) was subcloned into the kanamycin resistant pET28(a+) vector and the resultant plasmid transformed into BL21 (DE3) *Escherichia coli* cells for expression. Expression and purification were performed similarly to previous methods,⁶ with some changes.⁵ Cells were grown in Luria Broth at 30°C to an OD₆₀₀ of 0.6, and protein expression was induced with 1mM IPTG for 5 hours at 25°C. Cell pellets were re-suspended in buffer (pH 7.9) containing 0.5M NaCl, 20mM Tris-HCl, 5mM imidazole, 0.1mM EDTA, 1mM DTT, 10mg/mL PMSF and sonicated on ice (Sonic Dismembrator 550, Fischer Scientific) with 30 s pulses separated by 30 s breaks for a total of 20 minutes.

The lysate was cleared by centrifugation at 14,000g for 30 minutes. The soluble sfGFP was bound and eluted from a Ni-NTA resin (Novagen) according to manufacturer's protocol. sfGFP was purified further on a Sephacryl S-40(Pharmacia Biotech) size-exclusion column to remove His-rich proline isomerase in 25mM NaPO₄ and 100mM NaCl at pH 6.5. Fractions containing sfGFP were analyzed by SDS-PAGE for purity and the purified protein

was collected and concentrated and stored in 25mM NaPO₄ and 100mM NaCl at pH 6.5. Yield was determined by a theoretical extinction coefficient ($\epsilon=31\,519\text{ M}^{-1}\text{cm}^{-1}$).

NMR of Trapped Species

The trapped sfGFP species was prepared as previously described,⁵ and is explained in Figure 1. Triple-resonance NMR experiments were performed at 315K on a Bruker Avance800 equipped with a triple-resonance triple-axis gradient probe. Homonuclear TOCSY experiments were performed on a Varian 800MHz with triple-resonance triple-axis gradient probe at 315K with a 20ppm spectral width. Final conditions for each sample were 30 mM phosphate buffer (pH 6.8), and 100 mM NaCl, 95% H₂O/5% ²H₂O. All data were processed using NMRPipe.⁴⁴

HD Exchange

Solvent amide exchange was performed as described previously by following the peak intensity change in the ¹H-¹⁵N HSQC over time⁴⁵. Native and trapped samples were exchanged via Quick-Spin buffer exchange column into a 100mM phosphate, 100mM NaCl in 99.9% D₂O, pD=7.2 (pH = 6.8). Time between start of buffer exchange and first HSQC is 15:30, and first HSQC is completed by 25 minutes. Data were fit in Matlab using in-house scripts. Intrinsic exchange rates were calculated by Sphere⁴⁶.

Acknowledgments

We would like to thank the members of the laboratory for discussions of the work. Work was funded by the Molecular Biophysics Training Grant GM 08326(BTA), LANL (GW/PAJ) and NIH grants GM 54038 (PAJ & JNO) and DK 54441(PAJ).

References

1. Tsien RY. The green fluorescent protein. *Annu Rev Biochem.* 1998; 67:509–44. [PubMed: 9759496]
2. Barondeau DP, Putnam CD, Kassmann CJ, Tainer JA, Getzoff ED. Mechanism and energetics of green fluorescent protein chromophore synthesis revealed by trapped intermediate structures. *Proc Natl Acad Sci U S A.* 2003; 100:12111–6. [PubMed: 14523232]
3. Lu C, Bentley WE, Rao G. A high-throughput approach to promoter study using green fluorescent protein. *Biotechnol Prog.* 2004; 20:1634–40. [PubMed: 15575693]
4. Huang JR, Craggs TD, Christodoulou J, Jackson SE. Stable intermediate states and high energy barriers in the unfolding of GFP. *J Mol Biol.* 2007
5. Andrews BT, Schoenfish AR, Roy M, Waldo G, Jennings PA. The rough energy landscape of superfolder GFP is linked to the chromophore. *J Mol Biol.* 2007; 373:476–90. [PubMed: 17822714]
6. Pedelacq JD, Cabantous S, Tran T, Terwilliger TC, Waldo GS. Engineering and characterization of a superfolder green fluorescent protein. *Nat Biotechnol.* 2006; 24:79–88. [PubMed: 16369541]
7. Barondeau DP, Kassmann CJ, Tainer JA, Getzoff ED. Understanding GFP chromophore biosynthesis: controlling backbone cyclization and modifying post-translational chemistry. *Biochemistry.* 2005; 44:1960–70. [PubMed: 15697221]
8. Barondeau DP, Tainer JA, Getzoff ED. Structural evidence for an enolate intermediate in GFP fluorophore biosynthesis. *Journal of the American Chemical Society.* 2006; 128:3166–3168. [PubMed: 16522096]
9. Rosenow MA, Huffman HA, Phail ME, Wachter RM. The crystal structure of the Y66L variant of green fluorescent protein supports a cyclization-oxidation-dehydration mechanism for chromophore maturation. *Biochemistry.* 2004; 43:4464–4472. [PubMed: 15078092]
10. Rosenow MA, Patel HN, Wachter RM. Oxidative chemistry in the GFP active site leads to covalent cross-linking of a modified leucine side chain with a histidine imidazole: implications for the mechanism of chromophore formation. *Biochemistry.* 2005; 44:8303–11. [PubMed: 15938620]

11. Zhang LP, Patel HN, Lappe JW, Wachter RM. Reaction progress of chromophore biogenesis in green fluorescent protein. *Journal of the American Chemical Society*. 2006; 128:4766–4772. [PubMed: 16594713]
12. Ormo M, Cubitt AB, Kallio K, Gross LA, Tsien RY, Remington SJ. Crystal structure of the *Aequorea victoria* green fluorescent protein. *Science*. 1996; 273:1392–5. [PubMed: 8703075]
13. Patnaik SS, Trohalaki S, Pachter R. Molecular modeling of green fluorescent protein: structural effects of chromophore deprotonation. *Biopolymers*. 2004; 75:441–52. [PubMed: 15497152]
14. Weber W, Helms V, McCammon JA, Langhoff PW. Shedding light on the dark and weakly fluorescent states of green fluorescent proteins. *Proceedings of the National Academy of Sciences of the United States of America*. 1999; 96:6177–6182. [PubMed: 10339561]
15. Genick UK, Soltis SM, Kuhn P, Canestrelli IL, Getzoff ED. Structure at 0.85 Å resolution of an early protein photocycle intermediate. *Nature*. 1998; 392:206–9. [PubMed: 9515969]
16. Thompson MJ, Bashford D, Noodleman L, Getzoff ED. Photoisomerization and proton transfer in photoactive yellow protein. *J Am Chem Soc*. 2003; 125:8186–94. [PubMed: 12837088]
17. Habuchi S, Cotlet M, Gensch T, Bednarz T, Haber-Pohlmeier S, Rozenski J, Dirix G, Michiels J, Vanderleyden J, Heberle J, De Schryver FC, Hofkens J. Evidence for the isomerization and decarboxylation in the photoconversion of the red fluorescent protein DsRed. *Journal of the American Chemical Society*. 2005; 127:8977–8984. [PubMed: 15969574]
18. Maddalo SL, Zimmer M. The role of the protein matrix in green fluorescent protein fluorescence. *Photochem Photobiol*. 2006; 82:367–72. [PubMed: 16613487]
19. He X, Bell AF, Tonge PJ. Ground state isomerization of a model green fluorescent protein chromophore. *Febs Letters*. 2003; 549:35–38. [PubMed: 12914920]
20. Nifosi R, Tozzini V. Cis-trans photoisomerization of the chromophore in the green fluorescent protein variant E(2)GFP: A molecular dynamics study. *Chemical Physics*. 2006; 323:358–368.
21. Voliani V, Bizzarri R, Nifosi R, Abbruzzetti S, Grandi E, Viappiani C, Beltram F. Cis-trans photoisomerization of fluorescent-protein chromophores. *J Phys Chem B*. 2008; 112:10714–22. [PubMed: 18671358]
22. Dong J, Abulwerdi F, Baldrige A, Kowalik J, Solntsev KM, Tolbert LM. Isomerization in Fluorescent Protein Chromophores Involves Addition/Elimination. *J Am Chem Soc*. 2008
23. Mizuno H, Mal TK, Wälchi M, Kikuchi A, Fukano T, Ando R, Jeyakanthan J, Taka J, Shiro Y, Ikura M, Miyawaki A. Light-dependent regulation of structural flexibility in a photochromic fluorescent protein. *Proc Natl Acad Sci U S A*. 2008; 105:9927–9932.
24. Quillin ML, Anstrom DM, Shu X, O'Leary S, Kallio K, Chudakov DM, Remington SJ. Kindling fluorescent protein from *Anemonia sulcata*: dark-state structure at 1.38 Å resolution. *Biochemistry*. 2005; 44:5774–87. [PubMed: 15823036]
25. Andrews BT, Gosavi S, Finke JM, Onuchic JN, Jennings PA. The dual-basin landscape in GFP folding. *Proceedings of the National Academy of Sciences of the United States of America*. 2008; 105:12283–12288. [PubMed: 18713871]
26. Cody CW, Prasher DC, Westler WM, Prendergast FG, Ward WW. Chemical structure of the hexapeptide chromophore of the *Aequorea* green-fluorescent protein. *Biochemistry*. 1993; 32:1212–8. [PubMed: 8448132]
27. Esposito AP, Schellenberg P, Parson WW, Reid PJ. Vibrational spectroscopy and mode assignments for an analog of the green fluorescent protein chromophore. *Journal of Molecular Structure*. 2001; 569:25–41.
28. Schellenberg P, Johnson E, Esposito AP, Reid PJ, Parson WW. Resonance Raman scattering by the green fluorescent protein and an analogue of its chromophore. *Journal of Physical Chemistry B*. 2001; 105:5316–5322.
29. Chen MC, Lambert CR, Urgitis JD, Zimmer M. Photoisomerization of green fluorescent protein and the dimensions of the chromophore cavity. *Chemical Physics*. 2001; 270:157–164.
30. Tozzini V, Nifosi R. Ab initio molecular dynamics of the green fluorescent protein (GFP) chromophore: An insight into the photoinduced dynamics of green fluorescent proteins. *Journal of Physical Chemistry B*. 2001; 105:5797–5803.
31. Hyeon C, Dima RI, Thirumalai D. Pathways and kinetic barriers in mechanical unfolding and refolding of RNA and proteins. *Structure*. 2006; 14:1633–45. [PubMed: 17098189]

32. Seifert MH, Ksiazek D, Azim MK, Smialowski P, Budisa N, Holak TA. Slow exchange in the chromophore of a green fluorescent protein variant. *J Am Chem Soc.* 2002; 124:7932–42. [PubMed: 12095337]
33. Hopf M, Gohring W, Ries A, Timpl R, Hohenester E. Crystal structure and mutational analysis of a perlecan-binding fragment of nidogen-1. *Nat Struct Biol.* 2001; 8:634–40. [PubMed: 11427896]
34. Demidov VV, Dokholyan NV, Witte-Hoffmann C, Chalasani P, Yiu HW, Ding F, Yu Y, Cantor CR, Broude NE. Fast complementation of split fluorescent protein triggered by DNA hybridization. *Proc Natl Acad Sci U S A.* 2006; 103:2052–6. [PubMed: 16461889]
35. Orte A, Craggs TD, White SS, Jackson SE, Klenerman D. Evidence of an intermediate and parallel pathways in protein unfolding from single-molecule fluorescence. *Journal of the American Chemical Society.* 2008; 130:7898–7907. [PubMed: 18507381]
36. Xie JB, Zhou JM. Trigger factor assisted folding of green fluorescent protein. *Biochemistry.* 2008; 47:348–357. [PubMed: 18067273]
37. Enoki S, Saeki K, Maki K, Kuwajima K. Acid denaturation and refolding of green fluorescent protein. *Biochemistry.* 2004; 43:14238–48. [PubMed: 15518574]
38. Seifert MH, Georgescu J, Ksiazek D, Smialowski P, Rehm T, Steipe B, Holak TA. Backbone dynamics of green fluorescent protein and the effect of histidine 148 substitution. *Biochemistry.* 2003; 42:2500–12. [PubMed: 12614144]
39. Capraro DT, Roy M, Onuchic JN, Jennings PA. Backtracking on the folding landscape of the beta-trefoil protein interleukin-1beta? *Proc Natl Acad Sci U S A.* 2008; 105:14844–8. [PubMed: 18806223]
40. Chavez LL, Gosavi S, Jennings PA, Onuchic JN. Multiple routes lead to the native state in the energy landscape of the beta-trefoil family. *Proc Natl Acad Sci U S A.* 2006; 103:10254–8. [PubMed: 16801558]
41. Ferreira DU, Hegler JA, Komives EA, Wolynes PG. Localizing frustration in native proteins and protein assemblies. *Proc Natl Acad Sci U S A.* 2007; 104:19819–24. [PubMed: 18077414]
42. Huang H, Melacini G. High-resolution protein hydration NMR experiments: Probing how protein surfaces interact with water and other non-covalent ligands. *Analytica Chimica Acta.* 2006; 564:1–9. [PubMed: 17723356]
43. Melacini G, Kaptein R, Boelens R. Editing of chemical exchange-relayed NOEs in NMR experiments for the observation of protein-water interactions. *Journal of Magnetic Resonance.* 1999; 136:214–218. [PubMed: 9986764]
44. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR.* 1995; 6:277–93. [PubMed: 8520220]
45. Roy M, Jennings PA. Real-time NMR kinetic studies provide global and residue-specific information on the non-cooperative unfolding of the beta-trefoil protein, interleukin-1beta. *J Mol Biol.* 2003; 328:693–703. [PubMed: 12706726]
46. Bai Y, Milne JS, Mayne L, Englander SW. Primary structure effects on peptide group hydrogen exchange. *Proteins.* 1993; 17:75–86. [PubMed: 8234246]
47. DeLano, WL. The PyMOL Molecular Graphics System 0.98 edit. DeLano Scientific; Palo Alto, CA: 2002.

Abbreviations

FP	Fluorescent Protein
GFP	Green fluorescent protein
sfGFP	superfolder GFP
{N_{iso}}	isomerized native state
HSQC	heteronuclear single quantum coherence
HT	hula-twist

NMR	nuclear magnetic resonance
PF	protection factor
TOCSY	total correlation spectroscopy

Figure 1. The folding of sfGFP exhibits hysteresis with a long-lived, trapped intermediate

The non-coincidence of the simulated folding equilibrium curves show hysteresis. The red circles show the preparation of the native sample by equilibration at 1.8 M Gdn-HCl for 96 hours. The trapped sample is initially unfolded to 6.5M Gdn, and then equilibrated to 1.8 M Gdn-HCl for 96 hours. Both samples are put under NMR conditions by 2× spin columns into 0M Gdn-HCl.

Figure 2ABC. The structure of sfGFP still allows a hula-twist isomerization

While the highly regular barrel (A) holds the rings of the chromophore (green) in a rigid position, proline residues (magenta) are linked to chromophore formation (B). The chromophore is proposed to isomerize in a concerted motion (C) called a hula-twist (HT) isomerization. This conserves the volume of the chromophore while potentially leading to a weakly or non-fluorescent state. Figure was rendered using PyMol⁴⁷ using the superfolder crystal structure (2B3P)⁶.

Figure 3. Structural heterogeneity maps to prolines previously linked to isomerization
 Strip-type plots showing peaks with structural heterogeneity are mapped displayed. These peaks are displayed in Figure 4 in orange, and are on the lids of GFP.

Figure 4. Changes in residue-specific rates of solvent exchange and of residue-specific NMR spectral shifts are mapped onto the native and trapped structures of GFP and localize to the lid Exchange is consistent with previous work showing a highly stable barrel in both native and trapped structures, with high exchange occurring in the lid. Differences between the native and trapped structures dominate in the lid using several probes, as labeled. Figure was rendered using PyMol⁴⁷ using the superfolder crystal structure (2B3P)⁶. Proline residues discussed in the text (magenta in Figure 1) are displayed as spheres.

Figure 5. Homonuclear 2-D ^1H - ^1H TOCSY of native GFP (red) overlaid with trapped GFP (blue) shows an overall identical chromophore environment with subtle structural changes within the aromatic region. The assigned peaks for the chromophore spin system are labeled.

Figure 6. Homonuclear 2-D 1H - 1H TOCSY of native GFP (red) overlaid with trapped GFP (blue)

shows structural changes in the aromatic region. Protons discussed are shown labeled on the chromophore (A). Residues consistent with the spin system of model chromophore for GFP show changes consistent with isomerization, namely, in H_β (B & C). The $H_\epsilon H_\beta$ system shows heterogeneity in the native state, and averaging in the trapped state, evidence of chromophore flexibility in the trapped state. See Figure 5 for the full aromatic region.