Published in final edited form as:
Int Rev Cell Mol Biol. 2013 ; 302: 221–278. doi:10.1016/B978-0-12-407699-0.00004-2.

Beta-Barrel Scaffold of Fluorescent Proteins: Folding, Stability and Role in Chromophore Formation

Olesya V. Stepanenko[*], **Olga V. Stepanenko**[*], **Irina M. Kuznetsova**[*], **Vladislav V. Verkhusha**[**,1], and **Konstantin K. Turoverov**[*,1]

[*]Institute of Cytology of Russian Academy of Sciences, St. Petersburg, Russia

[**]Department of Anatomy and Structural Biology and Gruss-Lipper Biophotonics Center, Albert Einstein College of Medicine, Bronx, NY, USA

Abstract

This review focuses on the current view of the interaction between the β-barrel scaffold of fluorescent proteins and their unique chromophore located in the internal helix. The chromophore originates from the polypeptide chain and its properties are influenced by the surrounding protein matrix of the β-barrel. On the other hand, it appears that a chromophore tightens the β-barrel scaffold and plays a crucial role in its stability. Furthermore, the presence of a mature chromophore causes hysteresis of protein unfolding and refolding. We survey studies measuring protein unfolding and refolding using traditional methods as well as new approaches, such as mechanical unfolding and reassembly of truncated fluorescent proteins. We also analyze models of fluorescent protein unfolding and refolding obtained through different approaches, and compare the results of protein folding *in vitro* to co-translational folding of a newly synthesized polypeptide chain.

1. INTRODUCTION

Fluorescent proteins (FPs) are a powerful tool for the bioimaging of single molecules, intact organelles, live cells, and whole organisms. Fluorescence microscopy has become an invaluable approach in the fields of biochemistry, biotechnology, and cell and developmental biology. The great advantage of FPs in comparison with synthetic dyes and quantum dots is that they can be genetically introduced into cells, tissues, or whole organisms. FPs can be used to mark whole cells, to label and visualize single protein molecules, and to monitor their dynamics and interactions with other proteins.

Being enclosed in a β-barrel scaffold, the chromophore of FPs represents a unique fluorescent probe that is introduced into a target object within its own microenvironment (Fig. 4.1). The fluorescent properties of such a label are not sensitive to the environment but genetic engineering can be used to construct the microenvironment artificially. The construction of the first mutant variant of the wild-type green fluorescence protein (GFP) from the jellyfish *Aequorea victoria* with improved properties was the beginning of the continued development of FP variants. The new variants allow the use of advanced techniques, and development of novel methods requires the design of new FP variants. A pallet of FPs ranging from blue to far-red with different quantum yields, fluorescence lifetimes and photochemical characteristics has been developed. Remarkably, though some

[1]Corresponding authors: vladislav.verkhusha@einstein.yu.edu; kkt@incras.ru.

of these variants have <25% amino acid identity with the wild-type protein, all of them have a β-barrel fold with the chromophore inside.

It is generally established that the existence of a unique FP chromophore relies on the barrel and that its properties depend on the protein matrix vicinity. Much less is known about the influence of the chromophore on the β-barrel. This chapter summarizes what is known about the relationship between the β-barrel scaffold and the unique chromophore that it contains. The first two parts provide an overview of the current knowledge about the effect of the β-barrel scaffold on the chromophore properties. First, we summarize a variety of chromophore structures found in FPs, and we describe the autocatalytic and light-induced chromophore formation and transformations and discuss the interaction of the chromophore with protein matrix of the β-barrel. Then, we describe the structure and fluorescent properties of the chromophore from *A. victoria* GFP, its genetically engineered variants, and FPs from other organisms.

The last two parts are devoted to the problems of FP folding and stability, and how they are influenced by the chromophore. Internal interactions play a crucial role in all β-barrel proteins and in FPs, in particular, identifying them as globular proteins. For this reason, we begin with a brief reminder of the fundamental principles of globular protein folding and a description of the first examinations of the stability of different FPs. These studies were complicated by FP high stability, as quasi-stationary dependences were obtained only after several days of protein incubation in a denaturant. Furthermore, some FPs are prone to aggregation. The development of cycle3-GFP and superfolder-GFP (sfGFP), which do not aggregate and for which the refolded protein has the same properties as the native protein, allowed the possibility of a careful systematic study of unfolding–refolding processes. In this work, we attempt to analyze the reasons for the discrepancies in the description of these processes presented in different papers. We also analyze what is known from the literature on hysteresis during FP unfolding and refolding processes and the role that chromophore plays in it.

We present studies performed both using traditional methods to examine protein unfolding–refolding as well as new approaches, such as mechanical unfolding and reassembly of truncated FPs. We analyze the comparability of the results obtained using the different methods. We also compare the results of FP folding in vitro with co-translational folding of a newly synthesized polypeptide chain.

2. CHROMOPHORE FORMATION AND TRANSFORMATIONS IN FLUORESCENT PROTEINS

The fluorescence of FPs and their engineered variants spans the spectral range from blue to far-red. The main determinant of the emission hue is the chemical nature of the chromophore housed within the β-barrel. The amino acids in the chromophore environment fulfill multiple functions. Some of them contribute to chromophore synthesis; the contacts of other amino acids with the chromophore are related to the further adjustment of FP spectroscopic features and structural stability.

2.1. Chromophore Structures Found in Fluorescent Proteins

The ability of FPs to fluoresce in the visible spectral region descends from the chromophore hidden in the β-barrel scaffold. The position of the FP on spectral scale is mostly determined by the chemical structure of the chromophore, i.e. the more extended the system of π-conjugated electrons, the more red-shifted emission. Chromophore maturation does not require the involvement of any cofactors or enzymatic components except for molecular oxygen. The only prerequisite for the initiation of chromophore maturation is correct protein

folding, which results in the bending of the central α-helix, exposing the chromophore-forming tripeptide and arranging the catalytic amino acids in a position that is favorable for chromophore synthesis. The chromophore is self-generated from an internal tripeptide, X65Tyr66Gly67, through a multistep reaction that includes cyclization, dehydration, oxidation and, in some cases, hydrolysis. The first step of this complex reaction, tripeptide cyclization and the subsequent proton abstraction from the α-carbon of Tyr66, is supposed to be promoted by Arg96 and Glu222 with Arg96 playing the role of an electrostatic catalyst and Glu222 acting as base catalyst (Sniegowski et al., 2005; Wood et al., 2005). These amino acids are highly conserved among FPs. In addition to the catalytic function of Glu222, inspection of the changes in the crystallographic structure of EGFP induced by excessive X-ray irradiation revealed a stabilizing role of Glu222 (Royant and Noirclerc-Savoye, 2011). X-ray-induced EGFP bleaching was shown to be related to Glu222 decarboxylation and the associated rearrangement of the hydrogen bond network. On the basis of these observations, it was proposed that Glu222 contributes the rigidity of the chromophore cavity, thus restricting chromophore flexibility and preventing it from nonradiative deactivation of the excited state. The other absolutely conserved amino acid is the chromophore-forming glycine residue located at position 67. Substitution of Gly67 with any other residue impairs chromophore synthesis. It is believed that glycine is the only residue at position 67 that allows the formation of a central α-helix with the required kinked conformation. In that conformation of the α-helix, the amide nitrogen of Gly67 is in close proximity to the carbonyl carbon of the residue at position 65 and can perform a nucleophilic attack. In all natural FPs, position 66 is occupied by a Tyr residue but it can be replaced with any aromatic amino acid, as has been shown for artificial FP variants. Indeed, a cyan-emitting variant of GFP contains Trp at position 66, and a blue-emitting variant of GFP has a His residue instead of Tyr at position 66 (Tsien, 1998). It was shown that chromophore formation takes place in FP variants bearing Ser, Leu or Gly at position 66 but the resulting structures do not fluoresce and they instead undergo further reactions, such as hydrolysis (Barondeau et al., 2006, 2007). These data indicate that Tyr66 provides the proper oxidative chemistry during chromophore maturation and prevents undesirable side reactions, such as backbone fragmentation and hydrolysis. Analysis of more than 200 FP structures available in the Protein Data Bank revealed three highly conserved glycine residues located at positions 31, 33 and 35 (Ong et al., 2011). Interestingly, these internal amino acids are not involved in chromophore maturation, and their functions remain unclear.

A variety of chromophore structures is found in FPs (Fig. 4.2). The GFP-like green emitting chromophore that was the first to be discovered consists of two aromatic rings, including a phenolic ring from Tyr66 and a five-membered heterocyclic structure (Fig. 4.2*a*, (Cody et al., 1993; Ormo et al., 1996)). These two aromatic cycles are incorporated in the chromophore system through the bridge between them. The bridge is oxidized to have a double bond and to accomplish the conjugation of π-electrons from both aromatic rings. Such an expanded system of π-conjugated electrons is able to absorb and emit visible light.

The red emitting DsRed-like (according to DsRed from *Discosoma* species where it was found for the first time) chromophore contains an additional desaturated Cα–N bond at the Gln65 residue that further extends the system of π-conjugated electrons and results in a red shift of the absorption and emission (Fig. 4.2*b*; Gross et al., 2000). An entirely different type of red emitting chromophore is presented by Kaede-like chromophore (from the Kaede derived from the stony coral *Trachyphyllia geoffroyi*). Kaede-like chromophore is observed in a set of FPs, including Kaede, EosFP, dendFP and others (Matz et al., 2006). These proteins bear chromophores composed of three aromatic rings where a GFP-like chromophore core is supplemented by an indole ring from the His65 residue (Fig. 4.2*c*; Mizuno et al., 2003).

In contrast to the DsRed-like chromophores, the blue-emitting chromophore of mTagBFP (Subach et al., 2008) and mTagBFP2 (Subach et al., 2011a) has a shorter π-conjugated system of a five-membered heterocyclic structure and an N-acylimine double bond between the Cα and N atoms of the Leu65 residue, while the phenolic ring of the Tyr66 residue is out of conjugation being nearly perpendicular to the rest of the chromophore (Fig. 4.2*d*; Subach et al., 2010c).

There are at least four derivatives of the DsRed-like chromophore. Three of them are three-ring π-systems. The yellow chromophore of zFP538 from *Zoanthus* species has an additional tetrahydropyridine ring derived from the Lys65 residue (Fig. 4.2*e*; Remington et al., 2005). The third dihydro-oxazole ring of the orange chromophore found in mOrange, KO and its mutant variants is generated from the Thr65 residue (Fig. 4.2*f*; Shu et al., 2006). It is believed that less effective conjugation of the π-electrons is responsible for the spectroscopic features of yellow/orange FPs. The chromophore in the far-red photoswitched form of PSmOrange (Subach et al., 2011b), a mutant variant of mOrange, is also a three-ring system, in which the third dihydrooxazole ring is further oxidized to have a C=O double bond instead of a hydroxyl group (Fig. 4.2*f*). The far-red emission of light-induced PSmOrange and its enhanced version PSmOrange2 (Subach et al., 2012) is attributed to a more efficient π-conjugation of the GFP-like core with alylimine and carbonyl groups that are involved in the dihydrooxazole ring. Acylimine functionality of the DsRed-like chromophore in the chromoprotein asCP from sea anemone *Anemonia sulcata* and its derivative KFP undergoes hydrolysis between the carbon atom of the residue at position 64 and the N1 atom of the Met65 residue, leading to chromophore fragmentation (Fig. 4.2*g*; Quillin et al., 2005; Yampolsky et al., 2005). As we can see, position 65 of the chromogenic tripeptide can be occupied by any amino acid, which affects the chromophore chemistry and results in diverse chromophore structures.

Chromophores can adopt both *cis*- and *trans*-configurations with non-planar *trans*-chromophores seen mostly in nonfluorescent CPs while the nearly planar *cis*- and *trans*-chromophores are characteristic of proteins with a high quantum yield of fluorescence. An analysis of FP structures available in the Protein Data Bank revealed that there is a low frequency of perfectly planar chromophores in FPs (Maddalo and Zimmer, 2006). It was proposed that the chromophore cavity of FPs is not complementary to a planar chromophore, and thus the protein matrix induces chromophore deformation, twisting the phenolic ring of the Tyr66 residue slightly around Cα=Cβ double bond of the bridge. The ethylenic bridge is supposed to prevent the chromophore from undergoing a more prominent deformation. Upon excitation of the chromophore, the π-conjugation of the bridge is reduced and the phenolic ring of the chromophore can rotate freely. In this case, the protein matrix does not allow the chromophore to gain the perpendicularly twisted conformation that is postulated to be the main pathway of nonradiative energy dissipation (Megley et al., 2009). Thus, the microenvironment of the chromophore should be rigid enough in FPs with high quantum yield. Some of the recently developed FPs with far-red emission have a slightly relaxed chromophore microenvironment (Section 2.3). As a result, the quantum yield of those proteins suffers, but their chromophores gain the ability to equilibrate with the polar environment of the protein matrix. Thus, the more pronounced Stokes shift in those proteins is achieved at the expense of their quantum yield (Abbyad et al., 2007).

2.2. Autocatalytic and Light-Induced Chromophore Formation and Transformations

The general scheme of the autocatalytic chromophore maturation is presented in Fig. 4.3. The chromophore posttranslational chemistry is triggered by protein folding that brings the nitrogen of Gly67 and the carbonyl carbon of the residue at position 65 in close proximity to each other (Fig. 4.3*a*). The distance between the amide nitrogen of Gly67 and the carbonyl carbon of X65 at the deformed path of the α-helix inside the β-barrel was calculated to be

less than the sum of their covalent radii (Lemay et al., 2008). Following the folding, cyclization and oxidation events lead to the formation of the imidazolone-containing product that absorbs at approximately 350 nm and does not emit (Fig. 4.3*b*). Dehydration and oxidation of the Cα–Cβ bond of the Tyr66 residue result in green GFP-like chromophore formation (Fig. 4.3*c*). In this proton abstraction, the catalytic role is played by Arg96 and Glu222. The formation of the red DsRed-like chromophore is more complex (Miyawaki et al., 2012; Subach and Verkhusha, 2012). It includes the accumulation of TagBFP-like blue intermediate first (Pletnev et al., 2010; Subach et al., 2009a, 2009c) (Fig. 4.3*e*). The mechanism of N-acylimine formation of the TagBFP-like chromophore involves cyclization followed by oxidation/dehydration or dehydration/oxidation steps with Glu222 as a base catalyst in the proton abstraction. The red chromophore synthesis requires the final oxidation of the Cα–Cβ bond of the Tyr66 residue, and catalytic functions in this process are suggested for the pairs of Glu222/Lys69, Glu222/Arg69 and Glu222/Arg203 in the case of mCherry (Subach et al., 2009a), FTs (Pletnev et al., 2010) and PAmCherrys (Subach et al., 2009a). The hydroxyl group of the phenolic ring in the chromophores of GFPs and red FPs (RFPs) can be in a protonated or deprotonated state depending on the features of the chromophore microenvironment (Fig. 4.3*c,d* and *f,g*). For example, hydrogen bond formation between Tyr66 and Thr203 stabilizes the anionic form of the green chromophore in GFP and its derivatives (Ehrig et al., 1995; Heim et al., 1994). Substitution of Thr203 with Ile in GFP derivatives, such as sapphire-GFP, results in a predominantly neutral form of the green chromophore because its anionic form cannot be solvated in the absence of the hydroxyl group of Thr203. The introduction of the carboxyl groups in the vicinity of the chromophore is a way to shift the equilibrium to the neutral form of the chromophore (Shi et al., 2007; Shu et al., 2007). The carboxyl groups have pK_a values that are lower than the pK_a of the Tyr66 side chain hydroxyl of the chromophore in the ground state, which encourages the carboxyl ionization and stabilization of the chromophore in the neutral form. The anionic form of the red and green chromophores usually has a high quantum yield when buried inside the protein globule, while their neutral forms emit virtually none.

The photophysics of the green and red chromophores in some GFPs and RFPs is strongly affected by excited state proton transfer (ESPT). This reaction was revealed for the first time for wild-type GFP (Brejc et al., 1997). The protein has an absorption spectrum with two peaks at 395 and 475 nm attributed to the neutral and deprotonated forms of the green chromophore. The light absorption by the neutral form of the green chromophore enforces the proton abstraction from the hydroxyl of the Tyr66 residue, resulting in an intermediate excited state that differs from the excited state of the anionic chromophore by the conformation of the nearest amino acids to the chromophore. Both the intermediate excited state and the anionic chromophore emit a green light with maxima at 508 and 504 nm. Proton transfer takes place via a proton wire that is composed in GFP of the hydrogen bonded Ser65 and Tyr66 residues of the chromophore and a water molecule and the Ser205 and Glu222 residues. Recently, the second excited state intermediate preceding the fully deprotonated intermediate was observed by combined time resolved mid-infrared and visible pump-dump-probe spectroscopy experiments (Di Donato et al., 2011). The intermediate is characterized by a partial protonation of Glu222 and a shift of the protons involved in a hydrogen bond network. As a result of proton transfer events, the excitation of GFP at the absorption band of the neutral chromophore leads to a large Stokes shift of more than a 100 nm.

It was demonstrated that the same mechanism underlies the large Stokes shift of yellow/red FPs but their proton wires differ from those of GFP. For example, the high-resolution structure of red mKeima indicates that its chromophore is hydrogen bonded to the Asp165 residue via Ser148 (Henderson et al., 2009b). In contrast to mKeima, the recently developed LSSmKate1 and LSSmKate2 (Piatkevich et al., 2010a) and LSSmOrange (Shcherbakova et

al., JACS, 2012) proteins do not exhibit an additional absorption from the anionic form of the chromophore. In LSSmKate1 and LSSmKate2, the proton is accepted by carboxyl group of Glu167 directly or by carboxyl group of Ser165 via Asp167. On the basis of these data, it was demonstrated that incorporation of the carboxyl groups in position 165 or 167 creates an ESPT pathway and thus induces a large Stokes shift in several conventional RFPs, including mNeptune, mCherry, mStrawberry, mOrange, and mKO (Piatkevich et al., 2010b). LSSmKates are good probes for two-photon microscopy as they can be efficiently excited with standard two-photon light sources. Their far-red emission in combination with two-photon excitation allows for deep-tissue intravital imaging (Piatkevich et al., 2010a).

In addition to general autocatalytic reactions where the choice between the competitive green and red branches of synthesis and its endpoint is determined by the peculiarities of the chromophore microenvironment, almost all of the transitions of the scheme shown in Fig. 4.3 can be engineered to be time-depended or light-inducible. For example, fluorescent timers developed on the basis of mCherry, such as fast-FT, medium-FT, and slow-FT (Subach et al., 2009c), have time-delays that vary in their extent in the blue-to-red conversion (E-F transition in Fig. 4.3). The analysis of the crystal structures of the conventional blue-emitting Blue102 mutant of medium-FT and fast-FT combined with site-directed mutagenesis revealed the amino acids accounting for the timing properties of FTs (Pletnev et al., 2010). Amino acids at positions 69 and 82 affect the oxidation speed of Cα–Cβ of Tyr66, with Arg69 and Tyr82 delaying the oxidation and with Lys69 and Leu82 accelerating it. The rate of blue intermediate formation is determined by the amino acid at position 216: Ala and Cys residues at this position slow the acylimine formation, while Ser216 makes it faster. The blue-to-red conversion in Blue102 is blocked by the *trans* chromophore conformation, which is unfavorable for further oxidation. This is due to the presence of a bulky Ile at position 145 instead of Ser145 in fast-FT. The time-delayed blue-to-red conversion of FTs can be used for the temporal resolution of different cellular events.

The irreversible dark-to-red photoconversion of PATagRFP under UV light results in the bright red fluorescent form of the protein, which absorbs at 562 nm and emits at 595 nm (Subach et al., 2010a). Light-induced red fluorescence acquisition is supposed to be a two-step process involving the absorption of two photons by different chromophore intermediates. Presumably, the steps correspond to the B-to-E and E-to-F transitions shown in Fig. 4.3. The PATagRFP demonstrated a good performance in the multicolor single particle tracking photoactivated localization microscopy (spt-PALM) imagining of living cells (Subach et al., 2010a).

The irreversible light-induced E-to-F transition (Fig. 4.3) is realized in the photoactivatable FPs, such as the PAmCherries (Subach et al., 2009b). According to the crystallographic and mass spectrometric data, PAmCherries in the “dark-state” contain a nonplanar chromophore that is identical to the blue chromophore of mTagBFP. The oxidation of the Tyr66 Cα–Cβ bond in PAmCherries following UV light illumination is accompanied by the decarboxylation of the Glu222 and release of a CO_2 molecule via a Kolbe-like radical reaction. Photoactivated PAmCherries exhibit a bright red fluorescence with absorbance and emission maxima at approximately 570 and 595 nm, respectively. These proteins, as well as the other photo-activatable FPs, are intended for application in super-resolution techniques based on PALM (Subach et al., 2009b).

A number of irreversible and reversible photo-induced reactions that are not shown in Fig. 4.3 occur in FPs. We would like to briefly consider these reactions. A group composed of PAGFP, PSCFP and PSCFP2 is able to undergo irreversible phototransformation to a green emitting form (Chudakov et al., 2004). These proteins before illumination with UV light have the GFP-like green chromophore stabilized in a neutral form. Dark-to-green

photoactivation of PAGFP as well as cyan-to-green photoconversion of PSCFP and PSCFP2, similar to PAmCherries, is concomitant with decarboxylation of the Glu222 carboxyl group, which enforces a rearrangement of the chromophore environment and the successive chromophore ionization (Henderson et al., 2009a).

The Kaede subfamily of FPs, which includes Kaede (Ando et al., 2002), EosFP and its variant mEos2 (Wiedenmann et al., 2004), DendFP (Pakhomov et al., 2004) and its engineered monomeric versions Dendra and Dendra2 (Gurskaya et al., 2006), mcavRFP (Kelmanson and Matz, 2003), rfloRFP (Labas et al., 2002), mIrisFP (Adam et al., 2008), and mKikGR (Tsutsui et al., 2005), undergoes green-to-red photoconversion. The proteins in the Kaede subfamily hold the GFP-like two-ring chromophore in the green form and the three-ring Kaede-like chromophore in the red form. The His65 providing the third ring to the chromophore is supposed to be indispensable for Kaede-like chromophore synthesis. The green-like chromophore intermediate in the Kaede subfamily under light absorption is subjected to a series of photochemical reactions leading to the conjugation of the His65 indole ring with the GFP-like core through the Cα–Cβ double bond in His65 (Wiedenmann et al., 2011). The final red-emitting Kaede proteins contain a fluorescent anionic chromophore equilibrated with the nonfluorescent neutral form.

A set of FPs including cyan mTFP0.7 (Ai et al., 2006), green Dronpa (Ando et al., 2004), red asFP595 and its KFP derivative (Chudakov et al., 2003), rsCherry and rsCherryRev (Stiel et al., 2008), the green and red forms of IrisFP (Adam et al., 2008), and red rsTagRFP (Subach et al., 2010b) are capable of reversible photoconversion between fluorescent and nonfluorescent states under light illumination at a wavelength specific for the forward and backward transitions. The fluorescent state is supposed to contain an anionic chromophore, while the nonfluorescent state bears its neutral form, which is flexible enough to dissipate the excitation energy. Following the light absorption, a series of events occur involving a light-induced *cis–trans* isomerization of the chromophore along with the associated structural rearrangements within the chromophore's pocket and alterations of the hydrogen-bond network that change the protonation status of the chromophore (Henderson et al., 2007; Schafer et al., 2008).

An unusual mechanism that has not been previously described underlies a light-driven reversible photoconversion of Dreiklang, a mutant variant of yellow Citrine (Brakemann et al., 2011). The protein maintains a GFP-like two-ring chromophore in the protonation–deprotonation equilibrium of the Tyr66 hydroxyl. Excitation at the absorption band of the anionic chromophore of Dreiklang with the maximum at 515 nm results in bright yellow fluorescence that peaks at 529 nm. Intense illumination at the absorption band of the neutral chromophore of Dreiklang with the maximum at 405 nm induces a switching to a nonfluorescent state that absorbs at approximately 340 nm. The reverse kindling of the protein is achieved by the illumination of the protein with light at 340 nm. An inspection of the structure of Dreiklang in its on- and off-states by X-ray diffraction analysis and electrospray ionization mass spectrometry experiments revealed the addition/elimination of a hydroxyl group donated by the water in the chromophore vicinity to its five-membered imidazolinone ring at the C=N bond. This chemical modification disrupts the conjugation of the imidazolinone ring with the phenolic ring of the chromophore, and the final structure absorbs in the UV region and does not fluoresce. Remarkably, the wavelength of florescence excitation of Dreiklang (approximately 515 nm) is decouple from that used for switching it off and on (405 and 365 nm), which allows for avoidance of the interlocking of a switching and fluorescence readout in microscopic experiments.

2.3. Interaction of Chromophore with Protein Matrix of β-Barrel

The interactions of the chromophore with amino acids of its microenvironment can have an additional impact on the excitation and/or emission spectra position. The well-known example of the batochromic shift induced by substitution outside the chromophore is the red shift by 20 nm of the yellow fluorescence of YFP, a yellow enhanced version of GFP, with respect to its precursor wild-type GFP (Ormo et al., 1996). The chemical structure of the YFP chromophore remains the same as that of GFP. A 20-nm shift of emission is generated by a Tyr residue introduced instead of Thr203 in the position located above the chromophore. The Tyr is supposed to be involved in the π–π stacking interaction with the phenolic ring of Tyr66 from the chromophore (Wachter et al., 1998).

This approach is realized in the natural yellow-emitting phiYFP from sea medusa *Phialidium* species (Pakhomov and Martynov, 2011). PhiYFP has the most red-shifted spectrum of fluorescence among the proteins bearing the GFP-like chromophore. The fluorescence of phiYFP is red shifted by 10 nm in comparison with YFP derived from GFP, which is indicative of additional contacts that govern the spectroscopic properties of phiYFP in addition to π–π stacking. Two complementary structural factors that contribute to the yellow fluorescence of phiYFP have been proposed. They are an excitation-induced protonation of the N_2 nitrogen in the imidazolinone ring of the chromophore that takes place through the hydrogen bond network connecting the N_2 nitrogen with a proton donor Glu222 via the Thr65 and the destabilization of the negative charge at the phenolic ring of Tyr66 due to the absence of hydrogen bond donors in the vicinity of the phenol hydroxyl.

The opposite behavior has been observed in cyan amFP486 from *Anemonia majano* (Henderson and Remington, 2005) and mTFP1, a monomeric version of FP from *Clavularia* (Ai et al., 2008), where a stabilization of the negative charge of the phenol in the chromophore results in the blue shift of both absorbance and emission relative to that of GFP. The stabilization is achieved through the interaction of the Tyr66 phenolic ring with the positively charged His199, which is properly oriented against the chromophore by a well-organized hydrogen bond wire.

A batochromic shift of the spectra in the far-red FPs, mNeptune and mPlum, is also not coherent with covalent modification of a DsRed-like chromophore. It is attributed to the interaction of acylimine oxygen of the chromophore with the hydrogen bond donor in the chromophore-bearing pocket. In mPlum, the carboxyl group of the Glu16 residue donates the hydrogen bond that is weak in the ground state, but the light absorption stimulates the reorientation of the Glu16 side chain with subsequent strengthening of the hydrogen bond (Abbyad et al., 2007; Shu et al., 2009). That is, the proposed mechanism for the red shift in the emission with no effect on the absorption spectrum position of mPlum. In mNeptune, a water molecule that occupies a free space created by substitution of Met41 with Gly forms a strong hydrogen bond in both the ground- and excited-states leading to the red shift in the excitation and emission spectra (Lin et al., 2009). Similar to mNeptune, this mechanism causes the red shift in the recently developed far-red proteins eqFP650 and eqFP670 (Shcherbo et al., 2010). These proteins also contain a water-filled cavity due to the substitution of Met44 with less bulky amino acids.

As we can see, the chromophore makes numerous contacts with the protein matrix which affect the photophysics of the chromophore and tune the color of the fluorescence. We suppose that noncovalent interactions of the chromophore with the β-barrel do influence the protein stability. It has been shown that FP mutants that are unable to form chromophores have decreased stability with respect to their chromophore-bearing counterparts.

Indeed, a mutant form of GFP-S65T with the substitution of Gly67, a highly conserved amino acid that is strongly required for chromophore formation, to Ala is two times less stable against guanidine hydrochloride denaturation (Kutrowska et al., 2007) compared with EGFP (Stepanenko et al., 2004), with the mid-point of the transition being approximately 1.2 and 2.3 M for GFP-S65T/Gly67Ala and EGFP, respectively. An analysis of GdnHCl-induced denaturation of two mutant forms of sfGFP defective for chromophore synthesis has been performed (Andrews et al., 2007). The first mutant contains the substitution of catalytic Arg96 to Ala. The second has the substitutions Met88Tyr and Tyr74Met in positions preceding Pro75 and Pro89, precluding the correct conformation of α-helix. Both mutants of sfGFP exhibit significantly lower resistance to GdnHCl compared to sfGFP with the midpoint of transition being approximately 1.3, 0.8 and 4.2 M for sfGFP/Arg96Ala, sfGFP/Met88Tyr/Tyr74Met and sfGFP, respectively (Andrews et al., 2007).

3. STRUCTURE OF FLUORESCENT PROTEINS AND THEIR UNIQUE PROPERTIES

All FPs share the same β-barrel fold that in addition to being vital for fluorescence acquisition, donates them formidable stability. Together with the structural similarity, all FPs have the same drawbacks, mainly a propensity for oligomerization and aggregation. Here, we consider the peculiarities of structural and supramolecular organization of FPs and the ways to overcome their limitations. We start with wild-type GFP from the jellyfish *A. victoria* whose main disadvantage is poor folding at temperatures exceeding the temperature of its natural environment of cold boreal waters.

3.1. *Aequorea victoria* GFP and its Genetically Engineered Variants

Wild-type GFP, discovered first by Shimomura, is a small 25-kDa protein of 238 amino acids (Shimomura, 2006). Its polypeptide chain adopts a β-barrel scaffold that is vital for green fluorescence acquisition (Fig. 4.1). The eleven-stranded β-barrel of GFP is flanked by lids on both sides (Ormo et al., 1996). The barrel encloses an α-helix that runs up the barrel axis. This central α-helix is deformed substantially by the chromophore in the middle section. The chromophore inside the protein globule is protected from the bulk solvent by the barrel and lids. This spatial pattern is highly conserved among all GFP-like proteins.

Interestingly, careful examination of the 266 structures of GFP-like proteins available in the Protein Data Bank showed low variability of the lid segments of the β-barrel, especially of a lid that is opposite to the N- and C-termini (Ong et al., 2011). The amino acids in the next positions are highly conserved, corresponding to residues 89, 91 and 196 on the side where the N- and C-termini are situated and residues 20, 23, 27, 53, 101, 102, 104, 127, 130, 134 and 136 of the other lid of the barrel. In GFP, positions 89 and 196 are occupied by proline amino acids. The peptide bond preceding Pro196 is a typical *trans*-bond, while the side chain of the residue preceding Pro89 is in the *cis*-conformation. This behavior is characteristic of all GFP-like proteins with few exceptions (Ong et al., 2011). Moreover, the mutation of residues preceding Pro89 and Pro196 in GFP results in ineffective chromophore maturation (Andrews et al., 2007). Apparently, the X88-Pro89 and X195-Pro196 patches are important for maintaining of the kink in the α-helix backbone that is required for chromophore synthesis. The function of the highly conserved lid, which is curiously disordered, remains to be discovered. There is an assumption that the lid could be involved in some protein–protein interaction. This makes sense if we take into account that disordered protein segments or completely disordered proteins often have numerous partners and that they gain more structure during interaction with these partners (Turoverov et al., 2010).

GFP has a complex two-peaked excitation spectrum with a major peak at 395 nm and a minor one at 475 nm (Tsien, 1998). These two bands have been attributed to absorption by two chemically distinct chromophore species, i.e. the neutral and anionic form of the green chromophore. Excitation of both the neutral and the anionic chromophore results in green fluorescence with the maximum at 508 nm as the neutral chromophore tends to ionize in the excited state. Direct emission from a neutral chromophore occurs with low probability giving a shoulder in the emission spectra at 475 nm. Excited-state reactions of the chromophore are mediated by a proton wire that is composed of Ser65, Tyr66, Ser205 and Glu222 residues and a bound water molecule in GFP. It should be noted that many amino acids with side chains placed inside the barrel of GFP are charged or polar. Additionally, there are numerous bound water molecules inside the barrel cage. These residues and water molecules connected through hydrogen bonds are involved in chromophore synthesis and determine the photophysical behavior of the chromophore.

The excitation spectrum of GFP is strongly influenced by environmental factors, such as pH, temperature and small concentrations of denaturants and ions. For example, solution alkalization up to pH values of 10–11 leads to a decrease of the excitation band at 395 nm with a concomitant increase of the excitation band at 475 nm (Tsien, 1998). At elevated temperatures, similar changes are observed. The influence of small concentrations of denaturants and ions will be considered thoroughly in the next section. Wild-type GFP tends to form dimers at increased concentrations, and this aggregation also results in excitation spectrum deviations. In this case, the excitation band at 395 nm increases at the expense of the excitation band at 475 nm (Tsien, 1998).

GFP folds properly at temperature lower the 25 °C, and it accumulates in insoluble aggregates when expressed at 37 °C (Tsien, 1998). This drawback stimulated the creation of mutant variants of GFP with enhanced folding properties. One of these optimized variants is GFPmut1, which is identical to the commercially available and widely used EGFP (enhanced GFP) (Cormack et al., 1996). The GFPmut1/EGFP protein contains a substitution of Phe64 to Leu and Ser65 to Thr. A single mutation, Phe-64Leu, improves protein folding at 37 °C and the mutation Ser65Thr makes GFPmut1/EGFP manifold brighter compared to the wild-type protein. Still, the folding efficiency of GFPmut1/EGFP is quite low. Only 20% of GFPmut1/EGFP is present in the soluble form at 37 °C. The other variant of GFP with a reduced propensity for aggregation is cycle3-GFP or GFPuv (Patterson et al., 1997). It has three amino acid substitutions: Phe99Ser, Met153Thr and Val163Ala. The mutations such as Phe64Leu and Val163Ala are the most often used “folding enhancing” mutations in GFP mutants.

Provided by Invitrogen, a mutant form of Emerald is considered to be one of the best GFP variants based on its brightness and folding features (Shaner et al., 2005). The total list of substitutions in Emerald includes Phe64Leu, Ser65Thr, Ser72Thr, Asn149Lys, Met153Thr, Ile167Thr, and His231Leu. Careful examination of the impact of these substitutions on the brightness of bacterial colonies expressing the mutant form of GFP revealed that the most important substitutions for acquisition of green fluorescence are Phe64Leu, Ser65Thr, Ser72Thr, Asn149Lys, and Ile167Thr (Teerawanichpan et al., 2007). The substitutions Met153Thr and His231Leu were neutral in this regard. The mutation Met153Thr found in cycle3-GFP is believed to diminish protein aggregation at elevated temperatures. On the other hand, in a yellow variant of GFP, the mutation Met153Thr together with Val163Ala and Ser175Gly substitutions have been shown to improve the kinetics of chromophore maturation (Rekas et al., 2002). According to these observations, a VisGreen variant has been created bearing the minimal set of mutations that are indispensable for achieving the maximal visible fluorescence (Teerawanichpan et al., 2007). They are the aforementioned

Phe64Leu, Ser65Thr, Ser72Thr, Asn149Lys, and Ile167Thr. This version performed well in plant and animal cells.

The combination of the “enhanced GFP” mutations and the “cycle3” mutations yielded a GFP+ variant, a protein that possesses bright fluorescence and is able to fold correctly at 37 °C (Scholz et al., 2000; Waldo et al., 1999). It should be noted that GFP+ is also known as “folding-reporter-GFP.” Indeed, the folding efficiency, and hence the fluorescence, of GFP+ when it is fused downstream to a cellular protein is strongly correlated with the folding status of the fusion partner (Waldo et al., 1999). GFP+ has been widely used for screening soluble proteins (Pedelacq et al., 2002; van den Berg et al., 2006; Yokoyama, 2003), discovering drugs for degenerative diseases, such as Alzheimer (Kim et al., 2006), and identifying chaperones.

The last set of improved proteins includes super-proteins, such as sfGFP and supercharged-GFP. In contrast to the folding-reporter, the fluorescence of sfGFP is insensitive to misfolding of the fusion partner (Pedelacq et al., 2006). This robustly folded version of GFP has been generated from folding-reporter-GFP by fusing it downstream to poorly folded ferritin and screening libraries for fluorescent colonies. The additional destabilizing burden in the form of insoluble protein has resulted in obtaining sfGFP that possesses an improved tolerance for circular permutation, increased resistance against chemical denaturants and improved folding kinetics. In addition to the folding-reporter mutations, sfGFP bears six extra mutations. They are Ser30Arg, Tyr39Asn, Asn105Thr, Tyr145Phe, Ile171Val and Ala206Val. The greatest impact on the folding features of sfGFP was attributed to Ser30Arg substitution. The arginine residue in this position was shown by crystallography to induce the formation of a five-membered intramolecular ionic network connecting the first, second, fifth and sixth β-strands of the barrel. The ionic network consists of Glu32 (second β-strand), Arg30 (second β-strand), Glu17 (first β-strand), Arg122 (sixth β-strand), and Glu115 (fifth β-strand) residues. Substitution of Tyr39 with Asn was shown to initiate the formation of an α-helix at the loop between the second and third β-strands of the barrel, and it may contribute to the stability of this region. Substitutions of Tyr145 to Phe and Ile171 to Val are supposed to reduce the formation of off-pathway intermediates that are prone to aggregate, while substitutions of Asn105 to Thr and Ala206 to Val presumably increase the solubility of native protein.

It is worth discussing the eCGP123 GFP, although it is not a mutant of GFP, but rather it is derived from Azami-GFP from the stony coral (Kiss et al., 2009). eCGP123 tolerates overnight incubation at 80 °C without considerable loss of green fluorescence. To produce this extremely stable FP, the introduction of “folding interference” domains was applied, as it was performed in the development of sfGFP. Here, heterologous loops were used as destabilizing insertions. The loops were sequentially placed into three β-turns of GFP; the construct was subjected to evolution and selection steps after each insertion. This allows a gradual increase of protein stability, while simultaneous insertion of all three loops would destroy the FP completely.

On the basis of sfGFP with an initial charge of −7, a set of proteins with the charge ranging from −25 to +48 has been created by substituting highly solvent-exposed amino acids with negatively (Glu or Asp) or positively (Lys or Arg) charged residues (Lawrence et al., 2007). These proteins, termed as supercharged GFPs, preserve fluorescence and structure that are identical to those of sfGFP. The key feature of supercharged GFPs is their dramatically enhanced solubility. They remain soluble even in conditions strongly favoring aggregation, i.e. under heating at 100 °C for 1 min or in the presence of 40% of 2,2,2-trifluoroethanol, a chemical that stimulates protein aggregation. Their precursor sfGFP precipitated substantially under the conditions tested.

3.2. Fluorescent Proteins from Other Organisms

GFP-like proteins are found in numerous marine organisms. Together with GFP, 11 GFP-like proteins have been identified in the class Hydrozoa, the phylum Cnidaria until recently. Most of the Hydrozoan proteins are green FPs; the exceptions are a yellow FP from hydromedusa *Phialidium* sp. (phi-YFP) and a chromoprotein from an unidentified anthomedusa (Shagin et al., 2004). PhiYFP contains the typical green GFP-like chromophore and, to achieve yellow fluorescence, it utilizes the same structural approach, as it was realized in artificial EYFP derived from GFP, i.e. **π**–**π** stacking of the phenolic ring of the chromophore and Tyr203 placed above it. Recent findings imply that the Hydrozoan FP family is more abundant than previously believed. The first example of a multicolored FP in a single Hydrozoan species has been described recently (Aglyamova et al., 2011). GFP homologs of three colors, cyan, green and yellow, have been found in the medusa life cycle stage of the *Obelia* bioluminescent system, while previous attempts to clone FPs from the colonial polyp stage of *Obelia* have failed. The spectral variability of *Obelia* FPs most likely arises from the microenvironment of the green GFP-like chromophore. All three proteins tend to form stable aggregates composed of up to 128 monomers. In *Obelia*, these proteins localize in subcellular granules containing the photoprotein and FP. This allows for speculation that *Obelia* FPs could be involved in the regeneration of the photoprotein following its oxidation in a bioluminescence reaction (Aglyamova et al., 2011). Recently, two amazing representatives of GFP homologs have been identified in the class Hydrozoa. They are multidomain proteins, namely two-domain green abeGFP from the siphonophore *Abylopsis eschscholtzii* and the four-domain orange-fluorescent Ember from an unidentified jellyfish (Hunt et al., 2012). The analysis of the spectroscopic features of Ember revealed that only one domain contains an orange-emitting chromophore, which is likely a DsRed-like red chromophore; the other three domains bear green-emitting chromophores. Still, the final fluorescence of Ember is orange-red as a result of effective resonance energy transfer from the green subunits to the red one. The exact functions of the multidomain arrangement of these proteins are obscure, though the authors suggest that it contributes to overall protein stability (Hunt et al., 2012).

FPs are also present in bilateral animals of phyla Arthropoda and Chordata. All the nine endogenous FPs discovered in copepods of two families, Pontellidae and Aetedae (phylum Arthropoda, subphylum Crustacea), are green FPs (Hunt et al., 2010; Masuda et al., 2006; Shagin et al., 2004). Copepod GFPs exhibit extremely high brightness and exceedingly fast chromophore formation. The highest quantum yield of 0.92 that is close to the theoretical maximum has been reported for GFPs from copepod *Pontella mimocerami* (Shagin et al., 2004). Genome analysis revealed the occurrence of GFP genes in three amphioxus species of the *Branchiostoma* genus (phylum Chordata, subphylum Chepalochordata (Li et al., 2009)), while mRNA encoding GFPs was isolated only from *Branchiostoma floridae* species (Deheyn et al., 2007). Lancelet *B. floridae* harbors as many as 16 GFP-like proteins (Bomati et al., 2009). Two of the GFP-like proteins from *B. floridae* are likely not to carry a mature chromophore as they contain a 65Gly-Tyr-Ala67 tripeptide instead of a 65Gly-Tyr-Gly67, which is found in the rest of *B. floridae* GFP-like proteins. The only one of the chromophore-forming *B. floridae* GFP-like proteins is a brightly fluorescent green FP; the others are either weakly fluorescent or nonfluorescent. In addition to the diverse spectral features, these GFP-like proteins have distinct expression patterns that suggest that they perform different functions in the animal (Bomati et al., 2009).

The most abundant and color diverse is the family of GFP-like proteins found in sea anemones and corals (phylum Cnidaria, class Anthozoa). Their spectra span the range from cyan to red and purple-blue, and nonfluorescent colors are found as well (Verkhusha and Lukyanov, 2004). This class of GFP homologs suggests two options for red color that are realized through DsRed-like and Kaede-like chromophores.

Virtually all GFP-like proteins are oligomeric (Fig. 4.1). Being monomeric at low protein concentration, GFP tends to form dimers at increased protein concentration; under physiological conditions, it exists as a heterotetramer with aequorin, whose excitation energy it accepts and re-emits in the green range of the spectrum (Tsien, 1998). The weak dimerization of GFP and related proteins is easily alleviated by mutation of one of the three hydrophobic residues on the protein surface to charged residues, resulting in charge repulsion (Ala206Lys, Leu221Lys, or Phe223Arg (Zacharias et al., 2002)). Substitution of Ala206 to a bulky Val residue in sfGFP also favors the monomeric form, sterically hindering dimerization (Pedelacq et al., 2006). The other Hydrozoa FPs are obligate dimers, as well as Chordata FPs (Bomati et al., 2009; Shagin et al., 2004). Nearly all Anthozoa and Arthropoda FPs generate tetrameric complexes even at nanomolar concentrations; they are also inclined to form aggregates (Hunt et al., 2010; Yanushevich et al., 2002). The strong tendency of Arthropoda FPs to aggregation is observed as visible protein precipitation during storage. The nature of aggregation is the same for both groups of proteins. It was shown that aggregation is facilitated by electrostatic interaction between positively charged N-terminal patches and negatively charged patches on the FP surface. The introduction of neutral or negatively charged residues instead of positively charged ones increases protein solubility significantly (Yanushevich et al., 2002).

The tetrameric molecule of GFP-like proteins looks like dimer of dimers, where each monomer interacts with two adjacent monomers through two interfaces that differ significantly both in their chemical composition and in the strength of their interactions (Evdokimov et al., 2006; Remington et al., 2005; Wiedenmann et al., 2005; Yarbrough et al., 2001). The weaker one is composed of a central hydrophobic cluster with a few polar amino acids around it. The disruption of this interface is easily achieved by mutating the hydrophobic residue at the dimeric interface to a positively charged arginine or lysine (Campbell et al., 2002). This junction is weakened substantially or completely in naturally occurring dimeric GFP homologs (Loening et al., 2007; Wilmann et al., 2005). The second interface that is present in both dimeric and tetrameric GFP homologs is more extended; it involves numerous hydrogen bonds and salt bridges between polar residues and buried water molecules. The further stabilization of the region arises from the clasp between the C-termini of neighboring subunits (Yarbrough et al., 2001). Likewise, for the hydrophobic interface, the disruption usually starts with the introduction of charge disturbance, but the additional elimination of existing contacts between the polar groups and the modification of C-terminal patches of protein are required (Campbell et al., 2002). We should note that the interface-forming amino acids of both junctions are highly specific, even for GFP-like proteins from the same origin, which prevents them from heterooligomerization (Stepanenko et al., 2008). The tetrameric organization of Anthozoa and Arthropoda FPs results in a slight deformation of protein subunits, although their total tertiary structure remains conserved (Yarbrough et al., 2001). Thus, the β-barrel of GFP-derived proteins is near perfectly circular in cross-section (Ormo et al., 1996), while the β-barrel of Anthozoa and Arthropoda FPs has an elliptical shape (Yarbrough et al., 2001). This deformation is deemed to be pertinent to chromophore maturation. This is demonstrated by the fact that the disruption of tetramers is often impossible without compromising the fluorescence (Campbell et al., 2002). More specifically, the disruption of the stronger hydrophilic interface always eliminates the fluorescence, while the disruption of the hydrophobic interface usually gives a slightly less or poorly fluorescent mutant form. Therefore, the easy task of engineering the monomeric versions of proteins for which crystallographic data are available is heavily complicated by the need to subject the proteins to further optimization procedures for recovery of the fluorescence. For example, to create a monomeric and fluorescent version of red FP from *Discosoma* coral, as many as 33 amino acid substitutions were introduced (Campbell et al., 2002). That notwithstanding, there are currently monomeric FPs of various colors from violet-blue to far-red (Chudakov et al., 2010).

4. PIONEERING STUDIES OF FLUORESCENT PROTEIN STABILITY

FPs belong to the proteins with β-barrel topology (Tsien, 1998). Crystallographic structures revealed that these proteins resemble an 11-stranded β-can wrapped around a single central helix, in the middle of which is the chromophore (Ormo et al., 1996; Wachter et al., 1998; Yang et al., 1996). The cylinder has a diameter of ~30 Å and a length of ~40 Å (Yang et al., 1996). In the Protein Data Bank, there are more than 1000 structures of proteins with β-barrel topology. Among the proteins with this structure, 68% are various types of enzymes, and the remainder is binding proteins and transport membrane proteins. Most of them have six or eight β-strands. Membrane proteins have 14–20 β-strands. Among the other proteins with β-barrel topology, only FPs have an 11 β-stranded barrel (Nagano et al., 1999). The extreme resistance to a variety of denaturing effects is a special feature of FPs. For example, they are approximately 1000 times more stable than proteins of the lipocalin superfamily (Stepanenko et al., 2012a). The structure of odorant-binding proteins (pOBP), which represents a β-barrel composed of eight β-strands with a central nonpolar cavity for the binding of hydrophobic odorant molecules, possesses stability that is close to that of the GFP-like proteins (the difference in the free energy of pOBP in the native and unfolded states in the absence of a denaturant is −25 kJ/mol). At the same time, the rate of unfolding of pOBP is 1000 times faster than that of the GFP-like proteins under the same conditions (Staiano et al., 2007). Interestingly, azurin, which has a single tryptophan residue (Trp48) deeply buried in the hydrophobic central cavity of the β-barrel and thus resembles the FPs, is also much less stable to denaturant action (Gabellieri et al., 2008).

Despite an increasing number of studies on the stability and unfolding–refolding process of proteins with β-barrel topology in recent years, these processes have been much less understood than similar processes in α- and α/β-proteins. The β-barrel proteins are typical globular proteins. Before discussing the features of green FP folding, we will briefly discuss the fundamental principles of globular protein folding, which have been developed based on more than half a century of intensive investigation of this problem (see, e.g. Finkelstein and Ptitsyn, 2002; Nolting, 1999).

4.1. Fundamental Principles of Globular Protein Folding

The work of Anfinsen represents the earliest investigation of globular protein folding (Anfinsen, 1973). This work showed that the three-dimensional native structure of each protein is determined by its amino acid sequence. The polypeptide chain apparently adopts the structure corresponding to the minimum of free energy. Somewhat later, it was shown that the primary structure defines not only the three-dimensional structure of the protein in its native state but also a way of achieving it and the existence of a sufficiently high energy barrier between the native and unfolded states. The last point is extremely important because the existence of a high energy barrier between the native and fully or partially unfolded states means the impossibility of the existence of partially folded native states of proteins. Globular protein can be either native or denatured, partially or fully unfolded. This determines the reliability of performance by the globular proteins of their functions. It is this fact that allows for the possibility of obtaining crystals of proteins and, consequently, the ability to determine their structure by X-ray analysis. In this way, globular proteins drastically differ from synthetic polymer molecules. At the end of the twentieth century, it became apparent that many proteins have an amino acid sequence that, in principle, does not allow them to fold up in a three-dimensional structure; these are the so-called “intrinsically disordered proteins” (IDPs). This term underlines the intrinsic amino acid sequence-determined inability of these proteins to form ordered structures. It is commonly thought that the smaller the content of hydrophobic amino acid residues and the higher the net charge of a polypeptide chain, the smaller the probability that this chain will fold into a compact globule, and, contrary, the greater the propensity of the IDPs to form complexes

with other proteins and self-aggregate if, for some reasons, they fail to interact with their partners. Due to these structural peculiarities, IDPs on the one hand play key roles in signaling, recognition, and regulation systems, but on the other hand, their aggregates are strongly related to many of the so-called conformational diseases and amyloidoses in particular. Studies of IDPs comprise a separate and intensively developing research field (see, e.g. Dunker et al., 2008; He et al., 2009; Turoverov et al., 2010; Uversky and Dunker, 2010). We will not stop here on these studies, especially because FPs represent a very different type of proteins, namely, they are globular proteins with many intramolecular contacts.

The main methodological approach for studying the folding and stability of globular proteins is the examination of their in vitro unfolding and refolding induced by external factors, such as changes in the denaturant concentration, pH and temperature of the solution. To this aim, one can record stationary and (or) kinetic dependencies of protein characteristics, the values of which differ in native and denatured states of protein.

If such dependences are determined for the parameters that are linearly related to the concentration of the protein (such as fluorescence intensity at a fixed wavelength of registration), then the difference in the free energies between the native and denaturated states can be evaluated on the basis of the stationary dependence of the fraction of protein in the native and unfolded states on the denaturant concentration, and the free energy barrier between the native and denaturated states can be characterized by the rate constant determined from measurement of the kinetic dependence of the change in this parameter. The use of intensive parameters (fluorescence spectrum position, the relation of fluorescence intensities at two wavelengths, or fluorescence polarization) is possible only for the qualitative characterization of conformational transitions. Some of these parameters, e.g. parameter that characterizes the fluorescence spectrum position, can also be used for determination of protein fractions in native and denaturated states after specially elaborated correction (Staiano et al., 2007).

According to the current view, protein folding is determined by the protein's energy landscape (Jahn and Radford, 2005; Radford, 2000). This landscape describes the dependence of the free energy on all the coordinates determining the protein conformation. The number of conformational states accessible by a polypeptide chain is reduced while approaching the native state. Therefore, this energetic surface is often called the "energy funnel." Under the influence of external denaturating factors, the energy landscape varies. Under native conditions, the native state of the protein, corresponding to a deep minimum of the free energy, is energetically favorable and, conversely, the unfolded state of the protein is energetically unfavorable, but it corresponds to the minimum of free energy at a high concentration of denaturant (Fig. 4.4). For a long time, it was thought that protein folding is similar to the crystallization process and that a protein can only exist in two states: native and unfolded, with nucleus formation being the limiting step in the folding process. This model, known as the "nucleation and growth" model, well describes the folding of small single-domain proteins that follow the "all-or-none" principle. At intermediate concentrations of denaturant, the free energy of the molecules in the native and unfolded state can be comparable. This means that under these conditions, the concentration of molecules in the native and unfolded states is slightly different. The change in the energy landscape with the increase in the denaturant concentration is shown in Fig. 4.4*a*.

To describe the folding of large proteins, the "sequential protein folding" model, also known as the "framework" or "hierarchic" model, was proposed (Ptitsyn, 1973). It suggests that folding starts with the backbone forming secondary structure elements, which then interact to form a more advanced folding intermediate; the specific packing of the side chains

concludes the process. Each stage of the folding process stabilizes the major structural elements formed in the previous state, suggesting the existence of several folding intermediates. In this case, the stationary curves can determine the population of the native, intermediate and unfolded states at each denaturant concentration, while several rate constants determined from the kinetic curves will characterize the existing energy barriers between the existing states.

The intermediate states, which have the structural elements of the native state, are called on-pathway states. For a long time, it was believed that there is one universal intermediate state referred to as the "molten globule state" (Ohgushi and Wada, 1983). Other partially folded intermediates (e.g. premolten globule and highly ordered molten globule) were later found (Uversky and Ptitsyn, 1996). Figure 4.4*b* shows the change in the energy landscape with the increase in the denaturant concentration for proteins with one on-pathway intermediate state.

Along with the on-pathway states, there can be off-pathway states (traps). These states have structural elements that are not met in the native state. Figure 4.4*c* shows that at reducing the denaturant concentration, both the on-pathway (I_{on}) and off-pathway (I_{off}) states can appear. Molecules that were caught in the "traps" can be folded only after a further decrease of denaturant concentration via unfolding the off-pathway elements.

In living cells, a newly synthesized protein finds itself in the "overcrowded" physiological cell medium, where the concentration of proteins, nucleic acids, and polysaccharides is as high as 400 mg/ml, and where macromolecules occupy up to 40% of the medium volume (Ellis, 2001; Zimmerman and Minton, 1993). Such conditions can greatly affect all biological processes, including protein folding, misfolding, and aggregation (Chebotareva et al., 2004; Minton, 2000; Uversky et al., 2002; van den Berg et al., 2000). The folding of proteins in the living cell is complicated by at least two factors: the existence of unfavorable contacts with "neighbors," and the appearance of the incorrect intramolecular contacts during a co-translational folding (Turoverov et al., 2010). Therefore, in order for the correct folding to occur, a set of special protein-helpers provides assistance. These are the chaperones and the enzymes that regulate the *cis–trans* isomerization of proline and the formation of the disulfide bridges. They prevent protein aggregation and misfolding, accelerate folding, and participate in protein transport (e.g. protein translocation through the membranes) (Bader and Bardwell, 2001; Fink, 1999; Gilbert, 1994; Schmid, 2001).

4.2. Comparative Studies of Green and Red Fluorescent Proteins

Studies of the stability and folding of FPs began almost simultaneously with the intensive studies aimed to create new mutants with improved properties. In the early studies of FPs, their reactions to different treatments were examined. In particular, the effects on wild-type *Aequorea* GFP of treatment with detergents (Bokman and Ward, 1981; Ward and Bokman, 1982), heating (Ward et al., 1982) and proteases (Chalfie et al., 1994) were studied.

The important breakthrough in the construction of new variants of GFP was the construction of EGFP, which has two amino acid replacements in comparison with wild type (Cormack et al., 1996). In particular, one of the most important substitutions was S65T, which shifted the equilibrium constant between the neutral and anionic forms of the chromophore toward the anionic form and, thus, significantly increased the fluorescence quantum yield. Surprisingly, despite the existence of numerous EGFP applications, its spatial structure was determined only recently (Royant and Noirclerc-Savoye, 2011).

The discovery of a distant homolog of GFP cloned from *Discosoma* coral (Matz et al., 1999), called DsRed for its significantly red-shifted excitation and emission maxima (558 and 583 nm, respectively), has attracted great interest and stimulated the appearance of

several papers on its structure, its stability to different denaturants and its processes of folding and unfolding (Verkhusha et al., 2003; Vrzheshch et al., 2000). The structure and stability of DsRed were studied in comparison with EGFP. Both proteins have similar β-barrel folds but possess different oligomeric organization and chromophore structures. These works represent a thorough examination of DsRed treated with different denaturing actions, including heating, GdnHCl, and changes in pH. A large variety of spectroscopic and fluorescence methods was used. The main conclusion was that DsRed is much more stable than EGFP and that the processes of unfolding are highly complex. The authors proposed a kinetic mechanism for DsRed denaturation that includes consecutive conversion of the initial state of the protein to the denatured state through three intermediates. The first intermediate still emits fluorescence, and the last one is subjected to irreversible inactivation. Because of tight DsRed tetramerization, it was suggested that the obligatory protonation of each monomer results in the fluorescence inactivation of the whole tetramer. The remarkable fluorescence stability of DsRed under all conditions that have been studied was attributed to a significant extent to its tetrameric organization.

This conclusion stimulated a systematic analysis of five FPs with different degrees of oligomerization (Stepanenko et al., 2004). For these proteins, the fluorescent and absorbance parameters, the near-UV and visible CD spectra, the accessibility of the chromophores and the tryptophans to acrylamide quenching, and the resistance of these proteins to the guanidine hydrochloride unfolding and the kinetics of the approaching the unfolding equilibrium were compared. In this paper, tetrameric zFP506 was shown to be dramatically more stable than the EGFP monomer, assuming that the association might contribute to the protein conformational stability. RFPs were shown to possess comparable conformational stabilities regardless of oligomerization: monomeric mRFP1 is the most stable species under the equilibrium conditions, and tetrameric DsRed1 possesses the slowest unfolding kinetics. EGFP was shown to be considerably less stable than mRFP1, whereas tetrameric zFP506 was the most stable species analyzed in this study. It was concluded that the quaternary structure, which is an important stabilizing factor, does not represent the only circumstance dictating the dramatic variations between the FPs in their conformational stabilities.

However, all these studies were complicated by the tendency of FPs to aggregate and by the extremely slow process of their unfolding. The study of FP unfolding–refolding became much more effective with the use of new FP variants with greatly reduced tendencies for aggregation.

In the course of searching for brighter GFP variants using a DNA shuffling approach, a mutant that was 42 times more fluorescent than the wild type was identified (Crameri et al., 1996). This variant has three mutations (Phe99Ser, Met53Thr, Val163Ala) and was named cycle3-GFP, also referred to as GFPuv. It was suggested that more efficient folding and the higher yields of cycle3-GFP were the results of significantly reduced aggregation. The DNA shuffling method was used to improve cycle3-GFP folding. For this purpose, the protein was fused to a very poorly folding protein (H-subunit of ferritin) (Pedelacq et al., 2006). After four rounds of selection, sfGFP was identified. These two variants, cycle3-GFP and sfGFP, became the favorite subjects of FP stability and folding investigations.

5. UNFOLDING–REFOLDING OF FLUORESCENT PROTEINS

Virtually all recent work on the processes of FP unfolding–refolding has been performed with the cycle 3 (Enoki et al., 2006, 2004; Fukuda et al., 2000; Huang et al., 2007, 2008; Melnik et al., 2011a) or sfGFP (Andrews et al., 2009, 2007; Stepanenko et al., 2012b) proteins being subjected to different denaturing effects, including the chemical denaturants GdnHCl (Andrews et al., 2007; Fukuda et al., 2000; Huang et al., 2007) and GTC

(Stepanenko et al., 2012b), changes in the pH of the solution (Enoki et al., 2006, 2004), and changes in the ionic strength (Hsu et al., 2010).

Since FPs have a unique chromophore and a single tryptophan residue, Trp57, the fluorescence characteristics of which are sensitive to the structure of the protein, and nonradiative energy transfer from Trp to the chromophore in the neutral state exists in the native state of the protein, fluorescent methods have been used in practically all papers to monitor changes in the structure of the protein under different treatment conditions (Andrews et al., 2007; Enoki et al., 2006, 2004; Fukuda et al., 2000; Huang et al., 2008; Melnik et al., 2011a; Orte et al., 2008; Stepanenko et al., 2012b). In addition, to characterize the processes of unfolding-refolding, CD (Enoki et al., 2004; Huang et al., 2007), SAX (Enoki et al., 2006), single-molecule fluorescence (Orte et al., 2008), and single-molecule mechanical unfolding (Bornschlogl and Rief, 2011; Dietz et al., 2006; Dietz and Rief, 2004; Mickler et al., 2007) as well as theoretical approaches (Andrews et al., 2008; Reddy et al., 2012) were used.

One would think that such intense study of the same subject by different groups using different experimental approaches would provide a comprehensive view of the processes of the folding–unfolding of proteins. However, the investigators agree only on the fact that these proteins have a very high resistance to denaturing effects. Most believe that the protein has a complex energy landscape with a number of different intermediate states, high energy barriers, and multiple pathways (Andrews et al., 2008, 2007; Chirico et al., 2006; Enoki et al., 2006, 2004; Fukuda et al., 2000; Huang et al., 2007; Jackson et al., 2006; Mickler et al., 2007). In nearly every work, intermediate states in the pathway of FP unfolding–refolding were detected, though the number of intermediates differs from one to four, depending on the use of a denaturant action and registered parameters.

5.1. Intermediate States on Pathway of Fluorescent Protein Unfolding

Practically all developments related to the folding and stability of FPs are discussed in the thorough review by Hsu et al. (2009) with more or less details. The authors presented an unbiased survey of the papers on the theme and discussed a variety of the conclusions on the pathways of FP unfolding–refolding, but they did not demonstrate a preference for any one point of view. It would be of interesting to understand why the results obtained using different methods do not coincide and what is the real pathway of FP unfolding–refolding.

The most traditional approach of investigation of protein folding is the study of its unfolding–refolding processes induced by different concentrations of chemical denaturants. As FPs are highly stable against chemical denaturation, the use of urea is absolutely unacceptable, and in the majority of papers GdnHCl (Andrews et al., 2007; Fukuda et al., 2000; Hsu et al., 2010; Huang et al., 2007; Orte et al., 2008) was used as a denaturating agent. Recently, one study used an even stronger denaturant, namely GTC (Stepanenko et al., 2012b).

In the studies of protein unfolding–refolding processes, it is used to recording the denaturant-induced changes of the parameters which are sensitive to the changes of protein structure, such as the intensity of the intrinsic fluorescence, the CD signal and others, and the conclusions about the existence of intermediate states are drawn on the basis of the form of the curve. For GFP, it was shown that the dependence of the steady-state fluorescence intensity of the green chromophore correlates with the CD signal change, and can therefore be used to study the processes of FP unfolding–refolding (Fukuda et al., 2000). Unfortunately, in most of the papers, no experimental curves are available, and only calculated dependences of the fraction of molecules in different structural states on the denaturant concentration are given. In all the papers using GdnHCl as denaturant, it was

noted that the dependences of the data on the concentration of the denaturant are quasi-stationary and change over time (Andrews et al., 2007; Fukuda et al., 2000; Huang et al., 2007). More clearly, this was shown in the work carried out by the group of Jenning (Andrews et al., 2007). All these dependences (Andrews et al., 2007; Fukuda et al., 2000; Huang et al., 2007) have sigmoidal character and are well described by a two-state transition (Andrews et al., 2007; Fukuda et al., 2000; Huang et al., 2007). However, it was shown that a better approximation is obtained using a three-state approximation (see Fig. 3 in Andrews et al., 2007 and Fig. 7e in Huang et al., 2007). Although the differences between these two approximations are subtle, the three-state approximation was favored, and the conclusion on the existence of a native-like intermediate state was performed.

At the same time, the absorption spectrum of the FP was not recorded in any of the above papers, although it is known that FP absorption spectra are complex. For YFP, a significant sensitivity of the absorption spectra to anions, such as halides, nitrates and thiosulfate, was shown (Wachter and Remington, 1999). The existence of two forms of a yellow chromophore was shown with the change of pH (Hsu et al., 2010; Seward and Bagshaw, 2009). Nonetheless, even in the study of YFP unfolding–refolding, experimental data on chromophore fluorescence intensity were not corrected for the change of absorption at the wavelength of excitation, though such correction must be performed not only for YFP but also in the study of any FP. It was shown that the change in the chromophore absorption spectrum with the change of denaturant has a general character (Stepanenko et al., 2012b). This is caused not only by the change of pH but also by denaturing agents, such as GdnHCl and GTC. Hence, if we are to judge the conformational changes of the protein, we have to eliminate the fluorescence changes induced by changes in the absorption of the chromophore.

The visible absorption spectrum of sfGFP changes dramatically over all ranges of GTC concentrations (Fig. 4.5*a*). However, the most pronounced alterations in the visible absorption spectrum of sfGFP are recorded at concentrations of GTC up to approximately 0.7 M (Fig. 4.5*a*). Here, the visible absorption spectra of sfGFP demonstrated a pronounced drop in the absorption band at 485 nm, which corresponds to the anionic form of the chromophore, with a concomitant rise in the absorption band at 390 nm, which corresponds to the neutral chromophore. Further changes in the intensity of both the absorption maxima can be described as sigmoid curves. Some blue shift of the absorption band at 485 nm is observed at a GTC concentration exceeding 1.3 M. The optical density at approximately 425 nm remains unaltered in all range of GTC concentrations (Fig. 4.5*a*). The presence of such an isosbestic point in the visible absorption spectra of sfGFP indicates that only two types of molecules exist, namely sfGFP molecules with neutral and anionic chromophore forms, and the observed alterations of the visible absorption spectra are caused by changes in the ratio between them. The chromophore fluorescence intensity was corrected on the spectral effect related to the change of absorption at the excited wavelength (Fig. 4.5*b*).

A joint analysis of the changes in the fluorescence intensity of chromophore and tryptophan residue, the parameter A, and the change of the elution volume in response to changes in the denaturant concentration suggested that the structure of the protein varies slightly in the range of 0–0.1 M GTC, it then remains constant up to 0.9 M, and the unfolding of the protein occurs in the range of 0.9–1.7 M GTC. Therefore, the complex nature of the change of fluorescence intensity up to 0.9 M GTC is of purely spectroscopic character, while in the range from 0.9 to 1.7 M GTC, the dependence of the fluorescence intensity is determined by two factors: the continuing change in the optical density and the process of protein unfolding. The characteristics, such as the dependence of the corrected fluorescence intensity of the chromophore, fluorescence anisotropy and the parameter A, on the

denaturant concentration provide no evidence about the intermediate states in the denaturating pathway.

Nevertheless, in the literature, there are a number of papers reporting the existence of intermediate states in the way of unfolding. Many studies indirectly indicate the possibility of the existence of intermediate states in the process of protein denaturation. Recent studies show that the stability of FPs is not uniform through the scaffold cylinder (Melnik et al., 2011b, 2011c; Orte et al., 2008). In particular, this was shown in the work by Orte et al. (2008). In this work, cycle3-GFP was probed by measuring the hydrogen/deuterium (H/D) NMR exchange rates of more than 157 assigned amid protons that contain nearly two-thirds of the GFP amid protons. It was found that amide protons in β-strands 7, 8, 9 and 10 have, on average, higher exchange rates than the others in the β-barrel. Approximately 40 amide protons were found that do not undergo significant exchange, even after several months. It was concluded that most of these residues play an important role in stabilizing the structure of the protein and that they are clustered into a core region encompassing most of the β-strands, at least at one end of the barrel. It was shown that the majority of them are located in β-strands 1–6.

The existence of two regions with different structural stability was confirmed by differential scanning microcalorimetry (Melnik et al., 2011b, 2011c). The microcalorimetric analysis of the nonequilibrium melting of cycle3-GFP and its two mutants, Ile14Ala and Ile161Ala, revealed that the temperature-induced denaturation of this protein most likely occurs in three stages. The first and second stages involve melting of a smaller hydrophobic cluster formed around the residue Ile161, whereas a larger hydrophobic cluster (formed around the residues I14) is melted only at the last cycle3-GFP denaturation step or remains rather structured, even in the denatured state (Melnik et al., 2011b).

A complex energy landscape with at least two intermediate states was suggested in theoretical work (Reddy et al., 2012), though it is not clear what conditions (what denaturant concentration) were used for the calculation and how the results will depend on the denaturant concentration. The shape of the energy landscape depends on the concentration of the denaturant (Section 4.2). Consequently, the rate of unfolding will depend on the denaturant concentration. The kinetic experiments performed using manual mixing and stopped flow exhibit two relaxation phases at 7.0 M Gdn-HCl and higher and three phases at 6.5 M GdnHCl and lower (Andrews et al., 2007).

Several papers have described five intermediate states of FP unfolding (Enoki et al., 2006). Several approaches were used in these examinations of cycle3-GFP unfolding–refolding, but the main conclusion that there are several intermediate states was made on the basis of intrinsic tryptophan fluorescence (Enoki et al., 2006). In these papers, protein denaturation was induced by changing the pH of the solutions. The recorded complex dependence of fluorescence intensity on the pH was interpreted as the existence of several intermediate states. At the same time, it was not taken into account that energy transfer from the tryptophan residue to the chromophore depends not only on the distance between them but also on the chromophore ionization state, which changes with pH. Furthermore, the change in the ionization of groups in the microenvironment of the tryptophan residues that are known (White, 1959) to induce fluorescence quenching was not taken into account. In cycle3-GFP, there is carboxyl group from Asp 216 in the vicinity of the indole ring of the tryptophan residue. Its ionization can be changed by reducing the solution pH, even before any structural change occurs. In total, GFP contains 18 Asp and 16 Glu residues whose ionization and remoteness from the indole ring could be changed by changing the solution pH. Additionally, the fluorescence of a tryptophan residue exposed to solution is quenched by water molecules. None of these factors were discussed, and apparently, they were not

taken into account in the papers by Enoki et al. (2006, 2004). By the way, the dependence of indole fluorescence on pH below 2.9 is quite well known (White, 1959).

Finally, single-molecule methods based on nonradiative energy transfer (Orte et al., 2008) and the mechanical unfolding (Bornschlogl and Rief, 2011; Dietz et al., 2006; Dietz and Rief, 2004; Mickler et al., 2007) were used to investigate FPs. The advantage of these methods is that they do not need ensemble averaging, and that is why they enable the identification of the existence of parallel unfolding pathways and intermediate states that may not be highly populated (Haustein and Schwille, 2004; Tinnefeld and Sauer, 2005). The unfolding of individual molecules by chemical denaturants is monitored by the changes in the Forster resonance energy transfer (FRET). Usually, two fluorescent dyes (donor and acceptor) must be attached to the target protein, but for FP, it was necessary to bind only one chromophore (Alexa647) because the yellow chromophore of Citrine works as a donor. It is noteworthy that, except for the use of traditional FRET, the authors used a more sophisticated methodology, two-color coincidence detection (TCCD), which makes use of simultaneous excitation of the donor and acceptor by two overlapped lasers (Li et al., 2003; Orte et al., 2006). This method has the advantage of being sensitive to FRET changes, and it allows the additional detection of fully unfolded proteins, in which the fluorescence of the intrinsic fluorophore is quenched but where there is still a signal from the attached reference dye. Unexpectedly, for Citrine-Alexa-647 in native conditions, the presence of two structured states was detected, one of which has a high-FRET efficiency and the other has a lower FRET signal. The authors suggested that the low-FRET species is a partially structured state. At the same time, it is not clear why this state was referred to as a “partially structured intermediate state.” By the conventional meaning, the intermediate state is a state that appears on the pathway of protein denaturation, and the population of molecules in this state strongly depends on the concentration of the denaturing agent: this state does not exist in native or completely unfolded conditions. However, a “low-FRET” state was found in native conditions, and its population was practically independent of the GdnHCl concentration. Therefore, the scheme of Citrine unfolding given in the work of Orte et al. (2008) was scarcely supported by the experimental evidence. Furthermore, in this work, the authors also did not take into account the change in the absorption spectrum of the chromophore that is of great importance in the analysis of FRET.

A new powerful technique has become popular for the investigation of the structural stability of molecules, namely single-molecule force experiments, however, it is not evident whether the comparison of the results obtained by this method and those gained from the experiments with chemical denaturants or heating is reasonable. In reality, chemical denaturants and heating influence the protein as a whole; the action is applied simultaneously to all parts of protein, while in the mechanical experiments, the force is applied to local points of the protein.

Nonetheless, a series of single-molecule force experiments was successfully performed in the Rief laboratory on GFP (Dietz et al., 2006; Dietz and Rief, 2004; Mickler et al., 2007), and these studies even elaborate protocol-like instructions for investigators who are beginning to use atomic force microscopy to study mechanical protein unfolding or refolding (Bornschlogl and Rief, 2011). In the first work of this series, it was shown that the mechanical unfolding of GFP proceeds via two metastable intermediate states, which were connected with the detachment of a seven-residue N-terminal α-helix from the β-barrel. It was shown that detachment of this small α-helix completely destabilizes GFP thermodynamically even though the β-barrel is still intact and can bear a load. The second intermediate state was found to be a molecule lacking a full β-strand from the N- or C-terminus (Dietz and Rief, 2004). In the following work, it was shown that the GFP structural response depends upon the direction of the strain application. It was shown that the energy

landscape and the three-dimensional deformation response of functional protein structures can be systematically explored by changing the direction of force application (Mickler et al., 2007).

5.2. Hysteresis in Unfolding and Refolding of Fluorescent Proteins

The unique structure of FPs is their chromophore on the kinked α-helix running through the center of the β-can. The chromophore forms from an autocatalytic reaction of the backbone, involving cyclization, oxidation, and dehydration reactions (Cody et al., 1993; Ormo et al., 1996). Chromophore formation follows the construction of the β-barrel and helix kinking (Barondeau et al., 2003). Despite being surrounded by an 11-stranded β-barrel, the chromophore may isomerize in a hula-twist motion (Andrews et al., 2009). Furthermore, it turns out that the chromophore can be reached by the molecules of the solvent. The latter is proved by the change of its absorption spectrum by low concentrations of chelates just after their addition to the solution (Stepanenko et al., 2012b). This, in turn, causes the change of chromophore fluorescence and can be considered as the appearance of an intermediate state, though it is not a structural, but a spectroscopic effect, and it must be taken into account in all structural examinations of FPs, though it has not been done anywhere.

The formation and maturation of the chromophore in FP is a long-continued process because it requires the overcoming of a high energetic barrier. However, after its formation, the chromophore in its turn plays a crucial role in protein stability. This was most convincingly shown for sfGFP (Andrews et al., 2008, 2009, 2007). It was found that sfGFP unfolding is fully reversible as 100% of the chromophore signal of the native protein was recovered under strong refolding conditions. At the same time, the curve of the fraction of unfolded protein as a function of the final denaturant concentration for the unfolding transition shifts to the lower concentration of denaturant during the time of equilibration. The unfolding transition was found to be practically unchanged only between 96 and 192 h, indicating that a quasi-equilibrium has been achieved. During this period of time, only the unfolding transition curve moved to the smaller concentration of denaturant, while the position of the refolding transition curve did not change. Even after 192 h when the quasi-equilibrium was achieved, the unfolding and refolding transition curves did not coincide, indicating the existing of hysteresis.

It was suggested that extremely slow processes of unfolding and refolding of FP are related to the isomerization of proline residues (Andrews et al., 2009, 2007; Enoki et al., 2004), which most certainly play an important role in refolding, although the main role in the unfolding process belongs to the chromophore. This was first shown clearly and convincingly in the work by Andrews et al. (2007). In this work, the unfolding–refolding processes of FPs without a chromophore were investigated. For this purpose, the sfGFP/Arg96Ala mutant was chosen because this FP variant was discovered to slow chromophore formation from minutes to months (Wood et al., 2005). The equilibrium unfolding and refolding of sfGFP/Arg96Ala was monitored by the change in tryptophan fluorescence as a function of increased denaturant concentration. The equilibrium unfolding and refolding transitions of this protein were found to be superimposed and showed no evidence of hysteresis. The midpoint of the transition corresponds to 1.3 M GdnHCl, which is significantly lower than the midpoint of the sfGFP quasi-equilibrium unfolding transition (approximately 4.2 M GdnHCl) and even the sfGFP equilibrium refolding transition (approximately 1.8 M GdnHCl) (Andrews et al., 2007). This result was confirmed by the examination of another FP mutant without a chromophore. In the laboratory of prof. Jennings, it was revealed that mutation of the N-terminal residues in two X-Pro peptide sequences Tyr74Pro75 and Met88Pro89 to Tyr74Met/Met88Tyr also hindered chromophore formation. These two mutations map to the helical cap of the barrel, which is tightly pinned to the barrel in the sfGFP. It was shown that the equilibrium refolding and unfolding

transitions for this variant also eliminated the hysteresis. These data support the hypothesis that the hysteresis observed in unfolding–refolding is related to chromophore. Interestingly, the cooperativity of the folding–unfolding transitions for the double mutant has decreased from that observed for Arg96Ala with the midpoint at approximately 0.8 M GdnHCl (Andrews et al., 2007). Unfortunately, it was not proven that the β-barrel scaffold of the protein was not disturbed by these mutations.

In many papers, the slowness of the FP's processes of unfolding and refolding is associated with its proline residue isomerization (Andrews et al., 2008, 2009, 2007; Enoki et al., 2006, 2004; Hsu et al., 2009; Jackson et al., 2006). It is well known that proline *cis-trans* isomerization plays a key role in the rate-determining steps of protein folding (see, e.g. Levitt, 1981).

Proline residues are unique among natural amino acids because the cyclic side chain of proline prevents the rotation of an N–Cα bond, and the peptide backbone has no amide hydrogen for hydrogen bonding. This amino acid residue has a relatively high intrinsic probability of existing as the *cis* rather than the *trans* peptide isomer (Brandts et al., 1975; Grathwohl and Wuthrich, 1976), whereas for other amino acids, the probability is much smaller (less than 10^{-3}, see Ramachandran and Mitra, 1976).

There are 10 proline residues in FPs. The first of them, Pro13, is located in the beginning of the first β-strand, four proline residues (Pro187, Pro192, Pro196, Pro211) are located between β-strands, and another four are in the central α-helix (Pro54, Pro56, Pro58, Pro75). Due to these proline residues, this α-helix is rather kinked and in reality, represents several pieces of α-helix (Fig. 4.6). All these Pro residues are in the *trans* form, while one, Pro89, is in the *cis* form. This residue is located between the fourth β-strand and the central α-helix, dramatically changing the direction of the polypeptide chain. Evidently, it plays a crucial role in the packing of the α-helix inside the β-barrel. Importantly, this is provided by the unfavorable *cis* isomer. Surely, we cannot exclude the role of other proline residues in the cap of the barrel (Pro187, Pro192, Pro196) or proline residues that surely play important roles in kinking the α-helix backbone (Pro54, Pro56, Pro58, Pro75).

The retarding action of proline isomerization on the rate of FP protein refolding was shown in the unfolding–refolding experiment of cycle3-GFP in the presence of cyclophilin A (CycPA), a protein with peptidyl-proline isomerase activity (Andrews et al., 2007; Enoki et al., 2004). Later, this was proved in experiments on cycle3-GFP unfolding–refolding in the presence of *Escherichia coli* trigger factor (TF) (Xie and Zhou, 2008). TF is an efficient molecular chaperone in the catalysis of protein folding reactions that are rate-limited by the isomerization of prolyl bonds (Scholz et al., 1997). It forms a protective shield for nascent polypeptides at the ribosome exit tunnel and assists in the folding of most newly synthesized polypeptide chains (Hoffmann et al., 2006). The spontaneous and TF-assisted folding of GdnHCl denatured cycle3-GFP was probed by tryptophan fluorescence and chromophore fluorescence (Xie and Zhou, 2008). It was shown that in the presence of low concentrations of TF, the fast recovery of Trp57 fluorescence and the fast recovery and slow adjustment phases of the chromophore fluorescence are accelerated, suggesting that those processes are all coupled to proline isomerization. However, with increasing TF concentration, the rate constants for the rapid burial of Trp57 and for the acquisition and adjustment of chromophore fluorescence were decreased. These results were explained by competition between catalysis and binding by TF. Previously it was shown that TF-assisted protein folding requires repeated binding and release cycles between TF and the folding intermediates (Huang et al., 2000). The higher the concentration of TF, the greater the chance of recapture of substrate intermediates by TF. It was suggested that this binding effect can lead to the arrest of folding (Huang et al., 2002, 2000), so that GFPuv folding was

no longer limited by proline isomerization. Thus, it can be concluded that though the limiting-rate role of all the proline residues together in FPs is demonstrated, the role played by each of them is the task of future investigations.

5.3. Circular Permutation and Reassembly of Split-GFP

Many proteins being cleaved into two parts can noncovalently reassemble into a stable and a functional state (Carey et al., 2007). In some cases, fragments can reassemble spontaneously; in others, it occurred after association with other proteins to which fragments of target protein were attached (Kerppola, 2006; Kim et al., 2007; Michnick et al., 2007). The results obtained in the Boxer laboratory testify to the unique ability of fragments of GFP molecules to reassemble spontaneously, which seems to be due to a highly developed system of intramolecular interactions in this protein, which also lead to its high stability (Kent et al., 2008). These authors found an association between recombinant sfGFP (with replacement Thr65Ser), consisting of the first 10 β-strands (GFP 1–10), and a synthetic peptide with the same amino acid sequence as the last β-strand of GFP (GFP 11).

GFP 1–10 isolated from inclusion bodies in denaturing solution does not initially exhibit the absorption or fluorescence characteristics of the GFP chromophore. After the addition of fully synthetic GFP 11, chromophore maturation was observed. As a result, a protein with properties indistinguishable from the intact protein was reconstituted. The lack of structure in GFP 1–10 and the observed formation of the chromophore only after the addition of GFP 11 suggest that GFP 11 induces the precyclization structural constraints necessary for chromophore formation. Thus, the reconstitution of native absorption, fluorescence, and excited-state dynamics, including the deuterium isotope effect, suggested that strand 11 was in the correct orientation, allowing for the reassembly of the fully functional protein. Both the protonated (A) and deprotonated (B) forms of the chromophore were present in whole GFP because Ser at position 65 was introduced. These results demonstrate that the GFP barrel can be reassembled from pieces where one piece is readily prepared on a peptide synthesizer, thus any natural or unnatural amino acid can be introduced. This system is ideally suited for studying the reassembly of β-barrel structures with a built-in fluorescence reporter, and, by using circular permutation, it may prove possible to apply the same strategy to any strand of the β-barrel.

In subsequent papers (Do and Boxer, 2011; Kent and Boxer, 2011; Kent et al., 2009), a technique to obtain a fragment consisting of the first 10 β-strands was improved. Fragment GFP 1–10 was obtained by recombination: a trypsin cleavage site was specifically engineered into the normally trypsin-resistant GFP loop between the β-strands 10 and 11. After digestion of the loop, the GFP remains intact and spectrally indistinguishable from the uncut protein. The GFP 11 strand was removed by denaturation and then replaced by a fully synthetic strand with any desired sequence containing natural or unnatural amino acids. It occurs that only if GFP 1–10 was refolded from the denaturing solution into a solution containing an excess of GFP 11, the newly associated (GFP 1–10–GFP 11) protein has properties indistinguishable from native GFP.

It was found that if GFP 1–10 was refolded without GFP 11, the properties of GFP 1–10 were quite different from those of the intact and reassembled GFP (Fig. 4.7). Additionally, the fluorescence quantum yield of this refolded form was surprisingly only approximately five times less than the native, fully folded protein (Fig. 4.7). This was unexpected as the fluorescence from denatured GFP 1–10 (or denatured native GFP) is particularly low because the chromophore is in an unstructured environment, and nonradiative decay pathways lead rapidly to the ground state. The high fluorescence quantum yield of the chromophore and the decreased yet still existent secondary structure of GFP 1–10 observed

by circular dichroism suggest that GFP 1–10 has some residual structure. Furthermore, it was shown that refolded GFP 1–10 does not bind to GFP 11.

However, it was observed that refolded GFP 1–10 does rebind GFP 11 in the presence of light (initially observed in room light). It was proposed that refolded GFP 1–10 has the chromophore in the *trans* configuration (denoted *trans*-GFP 1–10), but light activation creates a photostationary state with the *cis* configuration of the chromophore (*cis*-GFP 1–10), and only *cis*-GFP 1–10 can bind GFP 11. It was noted that *cis* and *trans* are applied to the chromophores in simple solvent, while these may be twisted somewhat from their ideal geometry by constraints in the protein.

The strategy of creating any structures that do not have one of the β-strands or α-helix begins with the synthesis of full-length GFP with an introduction to the relevant segments of the polypeptide chain loops with sites that can be selectively digested with proteases and circular permutations of the C- and N-termini. Then, the specially introduced fragments of polypeptides are selectively cleaved by proteases, and the split-off fragment is removed under denaturing conditions using chromatography. The deletion is replaced with its synthetic analog with the desired amino acid substitutions. Finally, after transfer to native conditions, one will obtain the protein with the desired properties. Using this approach, the proteins with synthetic β-strand 10 (Do and Boxer, 2011), β-strand 11 (Kent and Boxer, 2011), and the central α-helix with the chromophore (Kent et al., 2009) were created. This method offers a new approach to the construction of fluorescent markers and is a perfect illustration of a crucial role of intermolecular interactions in the formation of FPs.

5.4. Co-translational Folding of Fluorescent Proteins

The material presented above indicates that significant progress in understanding FP unfolding–refolding processes in vitro has been achieved. However, until recently, nothing was known about the processes of these proteins folding in the cell. Meanwhile, this question is of particular interest in the case of FPs because the folding of newly synthesized chains must be substantially different from the refolding of the protein in vitro, as in the latter case, it occurs in the presence of an already formed chromophore. In addition, in the native FPs, as in any other protein and, especially, in proteins with β-barrel topology, the role of the contacts of residues that are significantly remote along the polypeptide chain is important (Fig. 4.6). In vivo, the polypeptide chain of a protein is synthesized in the peptidyl transferase center of the large ribosomal subunit and proceeds vectorially from the N- to the C-terminus. The nascent chain elongates within a narrow tunnel of the ribosome, where the folding into a native structure is not possible, and at the output of the channel, it emerges into a crowded environment of cell cytoplasm. Therefore, it is usually assumed that after synthesis on the ribosome, the nascent polypeptide chain of the protein folds with the help of chaperones and chaperonins.

The co-translational folding of FPs was studied by examining the de novo folding yield of cycle3-GFP expressed by polyribosomes from *E. coli* cells (Ugrinov and Clark, 2010). It was shown that even the longest GFP nascent chain cannot fold to a native conformation, while the C-terminal residues are conformationally constrained within the ribosomal exit tunnel. At the same time, a GFP variant with a C-terminal extension (CFPex) to span the ribosome exit tunnel, placing all GFP residues outside the tunnel, exhibited measurable quantities of GFP fluorescence tethered to the ribosome (Ugrinov and Clark, 2010).

This result of the in vivo experiment was confirmed and extended by the examination of the folding of ribosome-attached nascent FPs that were synthesized in vitro from truncated RNA transcripts (Kelkar et al., 2012). To generate ribosome-bound FP intermediates of GFP (CFP, EGFP, Venus, and Citrine) and RFP (mStrawberry, mCherry, DsRed, and

mTangerine), the variants were fused to a C-terminal reporter protein (CFTR). The RNA transcripts were truncated downstream of the last FP codon and translated in RRL to capture different lengths of the FP C-terminus within the ribosome exit tunnel (Fig. 4.8). In contrast to in vivo expression where nascent chains of different lengths were isolated on polysomes (Ugrinov and Clark, 2010), the in vitro system contains transcripts of a given tether length (Fig. 4.8).

It was shown that the formation of the characteristic FP β-barrel is prevented by sequestration of only a few C-terminal residues within the ribosome exit tunnel. In contrast, folding proceeds unimpeded when the last C-terminus residue is extended at least 31 amino acids beyond the ribosome peptidyl transferase center. Thus, the ribosome constrains tertiary folding as expected, but it has no detectable influence on either the kinetics or yield once the C-terminus has exited from the tunnel. It was found that co-translational folding intermediates with 10 β-strands outside the exit tunnel remain kinetically trapped in a non-native, on-pathway intermediate structure that retains folding competence for prolonged periods of time. It was also shown that the final step in FP folding is relatively unaffected by the cellular folding environment. Kinetic analysis revealed that co-translational FP folding involves at least two steps: the formation of a partially folded intermediate and the slow incorporation of the eleventh β-strand (and possibly others) into the final barrel structure (Kelkar et al., 2012). Kelkar et al. (2012) proposed the scheme of co-translational folding of FPs. They proposed that folding occurs through a landscape characterized by rapid formation of a stable N-terminal folding intermediate that likely occurs coincidently with, and may be facilitated by vectorial elongation of the nascent chain. These events are followed by a slow, rate-limiting step after ribosome release that requires the eleventh β-strand to form the final barrel structure necessary for chromophore catalysis. In cells, these events would normally be coupled when synthesis is completed and the nascent chain is released from the ribosome. Interestingly, despite their conserved architecture, the folding rates for FPs are faster for genetically selected “superfolder” proteins. Future examination of these proteins could clarify the rate-limiting step and the co-translational folding intermediates.

6. CONCLUDING REMARKS

The β-barrel scaffold plays a crucial role in chromophore formation. A properly folded β-barrel is the essential prerequisite for the initiation of chromophore maturation. The most important features for chromophore synthesis are Arg96, which plays the role of an electrostatic catalyst, and Glu222, which acts as a base catalyst. These amino acids are absolutely conserved among FPs. The residue of Gly67 from the chromophore-forming tripeptide is also strongly required for the chromophore cyclization through nucleophilic attack of amide nitrogen of Gly67 to the carbonyl carbon of the residue at position 65. The residue of Tyr66 that is conserved in all natural FPs is believed to provide the proper oxidative chemistry during chromophore maturation and prevent undesirable side reactions, such as backbone fragmentation and hydrolysis. The most variable position is position 65 of the chromogenic tripeptide; the amino acid at this position affects chromophore chemistry and results in diverse chromophore structures. The extent of π-conjugation within the chromophore is the main determinant of the emission hue. However, numerous contacts of the chromophore with the protein matrix further affect the photophysics of FPs, including color adjustment and excited-state proton transfer reactions. Moreover, through these internal interactions, the chromophore tightens the protein structure. Indeed, FP proteins are extremely stable. The quasi-equilibrium dependence of protein unfolding is achieved only on the fourth day of its incubation in solutions of GdnHCl. The high stability of these proteins is clearly manifested when truncated GFP generated by removing some β-strand or even an internal helix is reassembled with a synthetic peptide corresponding to the truncated

fragment. Thus, the reassembled protein exhibits properties that are indistinguishable from the native one. The formidable stability that is characteristic to FPs greatly complicates the study of the processes of their unfolding–refolding. Nevertheless, due to the creation of improved GFP variants such as cycle3-GFP and sfGFP, these studies are now possible. As these proteins are not subjected to aggregation, their unfolding is reversible. However, renaturation occurs at much lower denaturant concentration than unfolding, so that unfolding and refolding dependences do not coincide. The main reason for the existence of hysteresis is the fact that the chromophore is already synthesized and must be correctly positioned when forming the β-barrel scaffold of protein. Obviously, an important role in this process is played by the prolines and in particular, by Pro89, which is the *cis* isoform. Despite the large number of studies on the folding process of FPs that have appeared recently, there is no consensus in the literature on the availability and number of intermediate states. The study of these processes is complicated by the heterogeneity of the protein in its native state due to the presence of protonated and anionic forms of the chromophore and the change in the relative content of these forms depending on the concentration of the denaturant. Despite the fact that the existence of two forms of the chromophore has been known for a long time, until recently, this effect has not been taken into account in the study of FP folding.

Acknowledgments

Our work reviewed in this chapter was supported by the Russian Federation MCB RAS and RFBP 12-04-31469-mol-a grants to K.K.T. and by the grants GM073913 and CA164468 from the US National Institutes of Health to V.V.V.

References

Abbyad P, Childs W, Shi X, Boxer SG. Dynamic Stokes shift in green fluorescent protein variants. Proc Natl Acad Sci USA. 2007; 104:20189–20194. [PubMed: 18077381]

Adam V, Lelimousin M, Boehme S, Desfonds G, Nienhaus K, Field MJ, Wiedenmann J, McSweeney S, Nienhaus GU, Bourgeois D. Structural characterization of IrisFP, an optical highlighter undergoing multiple photo-induced transformations. Proc Natl Acad Sci USA. 2008; 105:18343–18348. [PubMed: 19017808]

Aglyamova GV, Hunt ME, Modi CK, Matz MV. Multi-colored homologs of the green fluorescent protein from hydromedusa *Obelia* sp. Photochem Photobiol Sci. 2011; 10:1303–1309. [PubMed: 21614405]

Ai HW, Henderson JN, Remington SJ, Campbell RE. Directed evolution of a monomeric, bright and photostable version of *Clavularia* cyan fluorescent protein: structural characterization and applications in fluorescence imaging. Biochem J. 2006; 400:531–540. [PubMed: 16859491]

Ai HW, Olenych SG, Wong P, Davidson MW, Campbell RE. Hue-shifted monomeric variants of *Clavularia* cyan fluorescent protein: identification of the molecular determinants of color and applications in fluorescence imaging. BMC Biol. 2008; 6:13. [PubMed: 18325109]

Ando R, Hama H, Yamamoto-Hino M, Mizuno H, Miyawaki A. An optical marker based on the UV-induced green-to-red photoconversion of a fluorescent protein. Proc Natl Acad Sci USA. 2002; 99:12651–12656. [PubMed: 12271129]

Ando R, Mizuno H, Miyawaki A. Regulated fast nucleocytoplasmic shuttling observed by reversible protein highlighting. Science. 2004; 306:1370–1373. [PubMed: 15550670]

Andresen M, Wahl MC, Stiel AC, Grater F, Schafer LV, Trowitzsch S, Weber G, Eggeling C, Grubmuller H, Hell SW, Jakobs S. Structure and mechanism of the reversible photoswitch of a fluorescent protein. Proc Natl Acad Sci USA. 2005; 102:13070–13074. [PubMed: 16135569]

Andrews BT, Gosavi S, Finke JM, Onuchic JN, Jennings PA. The dual-basin landscape in GFP folding. Proc Natl Acad Sci USA. 2008; 105:12283–12288. [PubMed: 18713871]

Andrews BT, Roy M, Jennings PA. Chromophore packing leads to hysteresis in GFP. J Mol Biol. 2009; 392:218–227. [PubMed: 19577576]

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; 373:476–490. [PubMed: 17822714]

Anfinsen CB. Principles that govern the folding of protein chains. Science. 1973; 181:223–230. [PubMed: 4124164]

Bader MW, Bardwell JC. Catalysis of disulfide bond formation and isomerization in *Escherichia coli*. Adv Protein Chem. 2001; 59:283–301. [PubMed: 11868275]

Barondeau DP, Kassmann CJ, Tainer JA, Getzoff ED. Understanding GFP post-translational chemistry: structures of designed variants that achieve backbone fragmentation, hydrolysis, and decarboxylation. J Am Chem Soc. 2006; 128:4685–4693. [PubMed: 16594705]

Barondeau DP, Kassmann CJ, Tainer JA, Getzoff ED. The case of the missing ring: radical cleavage of a carbon-carbon bond and implications for GFP chromophore biosynthesis. J Am Chem Soc. 2007; 129:3118–3126. [PubMed: 17326633]

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 USA. 2003; 100:12111–12116. [PubMed: 14523232]

Bokman SH, Ward WW. Renaturation of Aequorea green-fluorescent protein. Biochem Biophys Res Commun. 1981; 101:1372–1380. [PubMed: 7306136]

Bomati EK, Manning G, Deheyn DD. Amphioxus encodes the largest known family of green fluorescent proteins, which have diversified into distinct functional classes. BMC Evol Biol. 2009; 9:77. [PubMed: 19379521]

Bornschlogl T, Rief M. Single-molecule protein unfolding and refolding using atomic force microscopy. Methods Mol Biol. 2011; 783:233–250. [PubMed: 21909892]

Brakemann T, Stiel AC, Weber G, Andresen M, Testa I, Grotjohann T, Leutenegger M, Plessmann U, Urlaub H, Eggeling C, Wahl MC, Hell SW, Jakobs S. A reversibly photoswitchable GFP-like protein with fluorescence excitation decoupled from switching. Nat Biotechnol. 2011; 29:942–947. [PubMed: 21909082]

Brandts JF, Halvorson HR, Brennan M. Consideration of the possibility that the slow step in protein denaturation reactions is due to cis-trans isomerism of proline residues. Biochemistry. 1975; 14:4953–4963. [PubMed: 241393]

Brejc K, Sixma TK, Kitts PA, Kain SR, Tsien RY, Ormo M, Remington SJ. Structural basis for dual excitation and photoisomerization of the *Aequorea victoria* green fluorescent protein. Proc Natl Acad Sci USA. 1997; 94:2306–2311. [PubMed: 9122190]

Campbell RE, Tour O, Palmer AE, Steinbach PA, Baird GS, Zacharias DA, Tsien RY. A monomeric red fluorescent protein. Proc Natl Acad Sci U S A. 2002; 99:7877–7882. [PubMed: 12060735]

Carey J, Lindman S, Bauer M, Linse S. Protein reconstitution and three-dimensional domain swapping: benefits and constraints of covalency. Protein Sci. 2007; 16:2317–2333. [PubMed: 17962398]

Chalfie M, Tu Y, Euskirchen G, Ward WW, Prasher DC. Green fluorescent protein as a marker for gene expression. Science. 1994; 263:802–805. [PubMed: 8303295]

Chebotareva NA, Andreeva IE, Makeeva VF, Livanova NB, Kurganov BI. Effect of molecular crowding on self-association of phosphorylase kinase and its interaction with phosphorylase b and glycogen. J Mol Recognit. 2004; 17:426–432. [PubMed: 15362101]

Chirico G, Cannone F, Diaspro A. Unfolding time distribution of GFP by single molecule fluorescence spectroscopy. Eur Biophys J. 2006; 35:663–674. [PubMed: 16786346]

Chudakov DM, Belousov VV, Zaraisky AG, Novoselov VV, Staroverov DB, Zorov DB, Lukyanov S, Lukyanov KA. Kindling fluorescent proteins for precise in vivo photolabeling. Nat Biotechnol. 2003; 21:191–194. [PubMed: 12524551]

Chudakov DM, Matz MV, Lukyanov S, Lukyanov KA. Fluorescent proteins and their applications in imaging living cells and tissues. Physiol Rev. 2010; 90:1103–1163. [PubMed: 20664080]

Chudakov DM, Verkhusha VV, Staroverov DB, Souslova EA, Lukyanov S, Lukyanov KA. Photoswitchable cyan fluorescent protein for protein tracking. Nat Biotechnol. 2004; 22:1435–1439. [PubMed: 15502815]

Cody CW, Prasher DC, Westler WM, Prendergast FG, Ward WW. Chemical structure of the hexapeptide chromophore of the *Aequorea* green-fluorescent protein. Biochemistry. 1993; 32:1212–1218. [PubMed: 8448132]

Cormack BP, Valdivia RH, Falkow S. FACS-optimized mutants of the green fluorescent protein (GFP). Gene. 1996; 173:33–38. [PubMed: 8707053]

Crameri A, Whitehorn EA, Tate E, Stemmer WP. Improved green fluorescent protein by molecular evolution using DNA shuffling. Nat Biotechnol. 1996; 14:315–319. [PubMed: 9630892]

Deheyn DD, Kubokawa K, McCarthy JK, Murakami A, Porrachia M, Rouse GW, Holland ND. Endogenous green fluorescent protein (GFP) in amphioxus. Biol Bull. 2007; 213:95–100. [PubMed: 17928516]

Di Donato M, van Wilderen LJ, Van Stokkum IH, Stuart TC, Kennis JT, Hellingwerf KJ, van Grondelle R, Groot ML. Proton transfer events in GFP. Phys Chem Chem Phys. 2011; 13:16295–16305. [PubMed: 21847481]

Dietz H, Berkemeier F, Bertz M, Rief M. Anisotropic deformation response of single protein molecules. Proc Natl Acad Sci USA. 2006; 103:12724–12728. [PubMed: 16908850]

Dietz H, Rief M. Exploring the energy landscape of GFP by single-molecule mechanical experiments. Proc Natl Acad Sci USA. 2004; 101:16192–16197. [PubMed: 15531635]

Do K, Boxer SG. Thermodynamics, kinetics, and photochemistry of beta-strand association and dissociation in a split-GFP system. J Am Chem Soc. 2011; 133:18078–18081. [PubMed: 21981121]

Dunker AK, Silman I, Uversky VN, Sussman JL. Function and structure of inherently disordered proteins. Curr Opin Struct Biol. 2008; 18:756–764. [PubMed: 18952168]

Dutta S, Burkhardt K, Young J, Swaminathan GJ, Matsuura T, Henrick K, Nakamura H, Berman HM. Data deposition and annotation at the worldwide protein data bank. Mol Biotechnol. 2009; 42:1–13. [PubMed: 19082769]

Ehrig T, O'Kane DJ, Prendergast FG. Green-fluorescent protein mutants with altered fluorescence excitation spectra. FEBS Lett. 1995; 367:163–166. [PubMed: 7796912]

Ellis RJ. Macromolecular crowding: an important but neglected aspect of the intra-cellular environment. Curr Opin Struct Biol. 2001; 11:114–119. [PubMed: 11179900]

Enoki S, Maki K, Inobe T, Takahashi K, Kamagata K, Oroguchi T, Nakatani H, Tomoyori K, Kuwajima K. The equilibrium unfolding intermediate observed at pH 4 and its relationship with the kinetic folding intermediates in green fluorescent protein. J Mol Biol. 2006; 361:969–982. [PubMed: 16889795]

Enoki S, Saeki K, Maki K, Kuwajima K. Acid denaturation and refolding of green fluorescent protein. Biochemistry. 2004; 43:14238–14248. [PubMed: 15518574]

Evdokimov AG, Pokross ME, Egorov NS, Zaraisky AG, Yampolsky IV, Merzlyak EM, Shkoporov AN, Sander I, Lukyanov KA, Chudakov DM. Structural basis for the fast maturation of Arthropoda green fluorescent protein. EMBO Rep. 2006; 7:1006–1012. [PubMed: 16936637]

Fink AL. Chaperone-mediated protein folding. Physiol Rev. 1999; 79:425–449. [PubMed: 10221986]

Finkelstein, AV.; Ptitsyn, OB. Protein Physics: a Course of Lectures. Academic Press; 2002.

Fukuda H, Arai M, Kuwajima K. Folding of green fluorescent protein and the cycle3 mutant. Biochemistry. 2000; 39:12025–12032. [PubMed: 11009617]

Gabellieri E, Balestreri E, Galli A, Cioni P. Cavity-creating mutations in Pseudomonas aeruginosa azurin: effects on protein dynamics and stability. Biophys J. 2008; 95:771–781. [PubMed: 18424505]

Gilbert HF. Protein chaperones and protein folding. Curr Opin Biotechnol. 1994; 5:534–539. [PubMed: 7765469]

Grathwohl C, Wuthrich K. NMR studies of the molecular conformations in the linear oligopeptides H-(L-Ala)n-L-Pro-OH. Biopolymers. 1976; 15:2043–2057. [PubMed: 963242]

Gross LA, Baird GS, Hoffman RC, Baldridge KK, Tsien RY. The structure of the chromophore within DsRed, a red fluorescent protein from coral. Proc Natl Acad Sci USA. 2000; 97:11990–11995. [PubMed: 11050230]

Gurskaya NG, Verkhusha VV, Shcheglov AS, Staroverov DB, Chepurnykh TV, Fradkov AF, Lukyanov S, Lukyanov KA. Engineering of a monomeric green-to-red photoactivatable fluorescent protein induced by blue light. Nat Biotechnol. 2006; 24:461–465. [PubMed: 16550175]

Haustein E, Schwille P. Single-molecule spectroscopic methods. Curr Opin Struct Biol. 2004; 14:531–540. [PubMed: 15465312]

Hayashi I, Mizuno H, Tong KI, Furuta T, Tanaka F, Yoshimura M, Miyawaki A, Ikura M. Crystallographic evidence for water-assisted photo-induced peptide cleavage in the stony coral fluorescent protein Kaede. J Mol Biol. 2007; 372:918–926. [PubMed: 17692334]

He B, Wang K, Liu Y, Xue B, Uversky VN, Dunker AK. Predicting intrinsic disorder in proteins: an overview. Cell Res. 2009; 19:929–949. [PubMed: 19597536]

Heim R, Prasher DC, Tsien RY. Wavelength mutations and posttranslational autoxidation of green fluorescent protein. Proc Natl Acad Sci USA. 1994; 91:12501–12504. [PubMed: 7809066]

Henderson JN, Ai HW, Campbell RE, Remington SJ. Structural basis for reversible photobleaching of a green fluorescent protein homologue. Proc Natl Acad Sci USA. 2007; 104:6672–6677. [PubMed: 17420458]

Henderson JN, Gepshtein R, Heenan JR, Kallio K, Huppert D, Remington SJ. Structure and mechanism of the photoactivatable green fluorescent protein. J Am Chem Soc. 2009a; 131:4176–4177. [PubMed: 19278226]

Henderson JN, Osborn MF, Koon N, Gepshtein R, Huppert D, Remington SJ. Excited state proton transfer in the red fluorescent protein mKeima. J Am Chem Soc. 2009b; 131:13212–13213. [PubMed: 19708654]

Henderson JN, Remington SJ. Crystal structures and mutational analysis of amFP486, a cyan fluorescent protein from *Anemonia majano*. Proc Natl Acad Sci USA. 2005; 102:12712–12717. [PubMed: 16120682]

Hoffmann A, Merz F, Rutkowska A, Zachmann-Brand B, Deuerling E, Bukau B. Trigger factor forms a protective shield for nascent polypeptides at the ribosome. J Biol Chem. 2006; 281:6539–6545. [PubMed: 16407311]

Hsu ST, Blaser G, Behrens C, Cabrita LD, Dobson CM, Jackson SE. Folding study of Venus reveals a strong ion dependence of its yellow fluorescence under mildly acidic conditions. J Biol Chem. 2010; 285:4859–4869. [PubMed: 19901033]

Hsu ST, Blaser G, Jackson SE. The folding, stability and conformational dynamics of beta-barrel fluorescent proteins. Chem Soc Rev. 2009; 38:2951–2965. [PubMed: 19771338]

Huang GC, Chen JJ, Liu CP, Zhou JM. Chaperone and antichaperone activities of trigger factor. Eur J Biochem. 2002; 269:4516–4523. [PubMed: 12230563]

Huang GC, Li ZY, Zhou JM, Fischer G. Assisted folding of D-glyceraldehyde-3-phosphate dehydrogenase by trigger factor. Protein Sci. 2000; 9:1254–1261. [PubMed: 10892818]

Huang JR, Craggs TD, Christodoulou J, Jackson SE. Stable intermediate states and high energy barriers in the unfolding of GFP. J Mol Biol. 2007; 370:356–371. [PubMed: 17512539]

Huang JR, Hsu ST, Christodoulou J, Jackson SE. The extremely slow-exchanging core and acid-denatured state of green fluorescent protein. HFSP J. 2008; 2:378–387. [PubMed: 19436495]

Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. J Mol Graph. 1996; 14(33–38):27–38.

Hunt ME, Modi CK, Aglyamova GV, Ravikant DV, Meyer E, Matz MV. Multi-domain GFP-like proteins from two species of marine hydrozoans. Photochem Photobiol Sci. 2012; 11:637–644. [PubMed: 22251928]

Hunt ME, Scherrer MP, Ferrari FD, Matz MV. Very bright green fluorescent proteins from the Pontellid copepod *Pontella mimocerami*. PLoS One. 2010; 5:e11517. [PubMed: 20644720]

Jackson SE, Craggs TD, Huang JR. Understanding the folding of GFP using biophysical techniques. Expert Rev Proteomics. 2006; 3:545–559. [PubMed: 17078767]

Jahn TR, Radford SE. The Yin and Yang of protein folding. FEBS J. 2005; 272:5962–5970. [PubMed: 16302961]

Kelkar DA, Khushoo A, Yang Z, Skach WR. Kinetic analysis of ribosome-bound fluorescent proteins reveals an early, stable, cotranslational folding intermediate. J Biol Chem. 2012; 287:2568–2578. [PubMed: 22128180]

Kelmanson IV, Matz MV. Molecular basis and evolutionary origins of color diversity in great star coral *Montastraea cavernosa* (Scleractinia: Faviida). Mol Biol Evol. 2003; 20:1125–1133. [PubMed: 12777529]

Kent KP, Boxer SG. Light-activated reassembly of split green fluorescent protein. J Am Chem Soc. 2011; 133:4046–4052. [PubMed: 21351768]

Kent KP, Childs W, Boxer SG. Deconstructing green fluorescent protein. J Am Chem Soc. 2008; 130:9664–9665. [PubMed: 18597452]

Kent KP, Oltrogge LM, Boxer SG. Synthetic control of green fluorescent protein. J Am Chem Soc. 2009; 131:15988–15989. [PubMed: 19839621]

Kerppola TK. Design and implementation of bimolecular fluorescence complementation (BiFC) assays for the visualization of protein interactions in living cells. Nat Protoc. 2006; 1:1278–1286. [PubMed: 17406412]

Kim SB, Otani Y, Umezawa Y, Tao H. Bioluminescent indicator for determining protein–protein interactions using intramolecular complementation of split click beetle luciferase. Anal Chem. 2007; 79:4820–4826. [PubMed: 17539598]

Kim W, Kim Y, Min J, Kim DJ, Chang YT, Hecht MH. A high-throughput screen for compounds that inhibit aggregation of the Alzheimer's peptide. ACS Chem Biol. 2006; 1:461–469. [PubMed: 17168524]

Kiss C, Temirov J, Chasteen L, Waldo GS, Bradbury AR. Directed evolution of an extremely stable fluorescent protein. Protein Eng Des Sel. 2009; 22:313–323. [PubMed: 19364809]

Kutrowska BW, Narczyk M, Buszko A, Bzowska A, Clark PL. Folding and unfolding of a non-fluorescent mutant of green fluorescent protein. J Phys Condens Matter. 2007; 19:285223. [PubMed: 20126640]

Kuznetsova IM, Sulatskaya AI, Povarova OI, Turoverov KK. Reevaluation of ANS binding to human and bovine serum albumins: key role of equilibrium microdialysis in ligand – receptor binding characterization. PLoS One. 2012; 7:e40845. [PubMed: 22829890]

Labas YA, Gurskaya NG, Yanushevich YG, Fradkov AF, Lukyanov KA, Lukyanov SA, Matz MV. Diversity and evolution of the green fluorescent protein family. Proc Natl Acad Sci USA. 2002; 99:4256–4261. [PubMed: 11929996]

Lawrence MS, Phillips KJ, Liu DR. Supercharging proteins can impart unusual resilience. J Am Chem Soc. 2007; 129:10110–10112. [PubMed: 17665911]

Lemay NP, Morgan AL, Archer EJ, Dickson LA, Megley CM, Zimmer M. The role of the tight-turn, broken hydrogen bonding, Glu222 and Arg96 in the post-translational green fluorescent protein chromophore formation. Chem Phys. 2008; 348:152–160. [PubMed: 19079566]

Levitt M. Effect of proline residues on protein folding. J Mol Biol. 1981; 145:251–263. [PubMed: 7265199]

Li G, Zhang QJ, Zhong J, Wang YQ. Evolutionary and functional diversity of green fluorescent proteins in cephalochordates. Gene. 2009; 446:41–49. [PubMed: 19615432]

Li H, Ying L, Green JJ, Balasubramanian S, Klenerman D. Ultrasensitive coincidence fluorescence detection of single DNA molecules. Anal Chem. 2003; 75:1664–1670. [PubMed: 12705600]

Lin MZ, McKeown MR, Ng HL, Aguilera TA, Shaner NC, Campbell RE, Adams SR, Gross LA, Ma W, Alber T, Tsien RY. Autofluorescent proteins with excitation in the optical window for intravital imaging in mammals. Chem Biol. 2009; 16:1169–1179. [PubMed: 19942140]

Loening AM, Fenn TD, Gambhir SS. Crystal structures of the luciferase and green fluorescent protein from *Renilla reniformis*. J Mol Biol. 2007; 374:1017–1028. [PubMed: 17980388]

Maddalo SL, Zimmer M. The role of the protein matrix in green fluorescent protein fluorescence. Photochem Photobiol. 2006; 82:367–372. [PubMed: 16613487]

Masuda H, Takenaka Y, Yamaguchi A, Nishikawa S, Mizuno H. A novel yellowish-green fluorescent protein from the marine copepod, *Chiridius poppei*, and its use as a reporter protein in HeLa cells. Gene. 2006; 372:18–25. [PubMed: 16481130]

Matz MV, Fradkov AF, Labas YA, Savitsky AP, Zaraisky AG, Markelov ML, Lukyanov SA. Fluorescent proteins from nonbioluminescent *Anthozoa* species. Nat Biotechnol. 1999; 17:969–973. [PubMed: 10504696]

Matz MV, Labas YA, Ugalde J. Evolution of function and color in GFP-like proteins. Methods Biochem Anal. 2006; 47:139–161. [PubMed: 16342417]

Megley CM, Dickson LA, Maddalo SL, Chandler GJ, Zimmer M. Photophysics and dihedral freedom of the chromophore in yellow, blue, and green fluorescent protein. J Phys Chem B. 2009; 113:302–308. [PubMed: 19067572]

Melnik BS, Molochkov NV, Prokhorov DA, Uversky VN, Kutyshenko VP. Molecular mechanisms of the anomalous thermal aggregation of green fluorescent protein. Biochim Biophys Acta. 2011a; 1814:1930–1939. [PubMed: 21816236]

Melnik TN, Povarnitsyna TV, Glukhov AS, Uversky VN, Melnik BS. Sequential melting of two hydrophobic clusters within the green fluorescent protein GFP-cycle3. Biochemistry. 2011b; 50:7735–7744. [PubMed: 21823681]

Melnik TN, Povarnitsyna TV, Solonenko H, Melnik BS. Studies of irreversible heat denaturation of green fluorescent protein by differential scanning microcalorimetry. Thermochim Acta. 2011c; 512:71–75.

Merritt EA, Bacon DJ. Raster3D: photorealistic molecular graphics. Methods Enzymol. 1997; 277:505–524. [PubMed: 18488322]

Michnick SW, Ear PH, Manderson EN, Remy I, Stefan E. Universal strategies in research and drug discovery based on protein-fragment complementation assays. Nat Rev Drug Discov. 2007; 6:569–582. [PubMed: 17599086]

Mickler M, Dima RI, Dietz H, Hyeon C, Thirumalai D, Rief M. Revealing the bifurcation in the unfolding pathways of GFP by using single-molecule experiments and simulations. Proc Natl Acad Sci USA. 2007; 104:20268–20273. [PubMed: 18079292]

Minton AP. Implications of macromolecular crowding for protein assembly. Curr Opin Struct Biol. 2000; 10:34–39. [PubMed: 10679465]

Miyawaki A, Shcherbakova DM, Verkhusha VV. Red fluorescent proteins: chromophore formation and cellular applications. Curr Opin Struct Biol. 2012; 22:679–688. [PubMed: 23000031]

Mizuno H, Mal TK, Tong KI, Ando R, Furuta T, Ikura M, Miyawaki A. Photo-induced peptide cleavage in the green-to-red conversion of a fluorescent protein. Mol Cell. 2003; 12:1051–1058. [PubMed: 14580354]

Nagano N, Hutchinson EG, Thornton JM. Barrel structures in proteins: automatic identification and classification including a sequence analysis of TIM barrels. Protein Sci. 1999; 8:2072–2084. [PubMed: 10548053]

Nolting, B. Biophysical Methods. Springer-Verlag; Berlin-Heidelberg: 1999. Protein Folding Kinetics.

Ohgushi M, Wada A. 'Molten-globule state': a compact form of globular proteins with mobile side-chains. FEBS Lett. 1983; 164:21–24. [PubMed: 6317443]

Ong WJ, Alvarez S, Leroux IE, Shahid RS, Samma AA, Peshkepija P, Morgan AL, Mulcahy S, Zimmer M. Function and structure of GFP-like proteins in the protein data bank. Mol Biosyst. 2011; 7:984–992. [PubMed: 21298165]

Ormo M, Cubitt AB, Kallio K, Gross LA, Tsien RY, Remington SJ. Crystal structure of the *Aequorea victoria* green fluorescent protein. Science. 1996; 273:1392–1395. [PubMed: 8703075]

Orte A, Clarke R, Balasubramanian S, Klenerman D. Determination of the fraction and stoichiometry of femtomolar levels of biomolecular complexes in an excess of monomer using single-molecule, two-color coincidence detection. Anal Chem. 2006; 78:7707–7715. [PubMed: 17105162]

Orte A, Craggs TD, White SS, Jackson SE, Klenerman D. Evidence of an intermediate and parallel pathways in protein unfolding from single-molecule fluorescence. J Am Chem Soc. 2008; 130:7898–7907. [PubMed: 18507381]

Pakhomov AA, Martynov VI. Probing the structural determinants of yellow fluorescence of a protein from *Phialidium* sp. Biochem Biophys Res Commun. 2011; 407:230–235. [PubMed: 21382348]

Pakhomov AA, Martynova NY, Gurskaya NG, Balashova TA, Martynov VI. Photoconversion of the chromophore of a fluorescent protein from *Dendronephthya* sp. Biochemistry (Mosc). 2004; 69:901–908. [PubMed: 15377271]

Patterson GH, Knobel SM, Sharif WD, Kain SR, Piston DW. Use of the green fluorescent protein and its mutants in quantitative fluorescence microscopy. Biophys J. 1997; 73:2782–2790. [PubMed: 9370472]

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]

Pedelacq JD, Piltch E, Liong EC, Berendzen J, Kim CY, Rho BS, Park MS, Terwilliger TC, Waldo GS. Engineering soluble proteins for structural genomics. Nat Biotechnol. 2002; 20:927–932. [PubMed: 12205510]

Piatkevich KD, Hulit J, Subach OM, Wu B, Abdulla A, Segall JE, Verkhusha VV. Monomeric red fluorescent proteins with a large Stokes shift. Proc Natl Acad Sci USA. 2010a; 107:5369–5374. [PubMed: 20212155]

Piatkevich KD, Malashkevich VN, Almo SC, Verkhusha VV. Engineering ESPT pathways based on structural analysis of LSSmKate red fluorescent proteins with large Stokes shift. J Am Chem Soc. 2010b; 132:10762–10770. [PubMed: 20681709]

Pletnev S, Subach FV, Dauter Z, Wlodawer A, Verkhusha VV. Understanding blue-to-red conversion in monomeric fluorescent timers and hydrolytic degradation of their chromophores. J Am Chem Soc. 2010; 132:2243–2253. [PubMed: 20121102]

Ptitsyn OB. Stages in the mechanism of self-organization of protein molecules. Dokl Akad Nauk SSSR. 1973; 210:1213–1215. [PubMed: 4721708]

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 A resolution. Biochemistry. 2005; 44:5774–5787. [PubMed: 15823036]

Radford SE. Protein folding: progress made and promises ahead. Trends Biochem Sci. 2000; 25:611–618. [PubMed: 11116188]

Ramachandran GN, Mitra AK. An explanation for the rare occurrence of cis peptide units in proteins and polypeptides. J Mol Biol. 1976; 107:85–92. [PubMed: 1003461]

Reddy G, Liu Z, Thirumalai D. Denaturant-dependent folding of GFP. Proc Natl Acad Sci USA. 2012; 109:17832–17838. [PubMed: 22778437]

Rekas A, Alattia JR, Nagai T, Miyawaki A, Ikura M. Crystal structure of venus, a yellow fluorescent protein with improved maturation and reduced environmental sensitivity. J Biol Chem. 2002; 277:50573–50578. [PubMed: 12370172]

Remington SJ, Wachter RM, Yarbrough DK, Branchaud B, Anderson DC, Kallio K, Lukyanov KA. zFP538, a yellow-fluorescent protein from *Zoanthus*, contains a novel three-ring chromophore. Biochemistry. 2005; 44:202–212. [PubMed: 15628861]

Royant A, Noirclerc-Savoye M. Stabilizing role of glutamic acid 222 in the structure of enhanced green fluorescent protein. J Struct Biol. 2011; 174:385–390. [PubMed: 21335090]

Schafer LV, Groenhof G, Boggio-Pasqua M, Robb MA, Grubmuller H. Chromophore protonation state controls photoswitching of the fluoroprotein asFP595. PLoS Comput Biol. 2008; 4:e1000034. [PubMed: 18369426]

Schmid FX. Prolyl isomerases. Adv Protein Chem. 2001; 59:243–282. [PubMed: 11868274]

Scholz C, Stoller G, Zarnt T, Fischer G, Schmid FX. Cooperation of enzymatic and chaperone functions of trigger factor in the catalysis of protein folding. EMBO J. 1997; 16:54–58. [PubMed: 9009267]

Scholz O, Thiel A, Hillen W, Niederweis M. Quantitative analysis of gene expression with an improved green fluorescent protein. p6. Eur J Biochem. 2000; 267:1565–1570. [PubMed: 10712585]

Seward HE, Bagshaw CR. The photochemistry of fluorescent proteins: implications for their biological applications. Chem Soc Rev. 2009; 38:2842–2851. [PubMed: 19771331]

Shagin DA, Barsova EV, Yanushevich YG, Fradkov AF, Lukyanov KA, Labas YA, Semenova TN, Ugalde JA, Meyers A, Nunez JM, Widder EA, Lukyanov SA, Matz MV. GFP-like proteins as ubiquitous metazoan superfamily: evolution of functional features and structural complexity. Mol Biol Evol. 2004; 21:841–850. [PubMed: 14963095]

Shaner NC, Steinbach PA, Tsien RY. A guide to choosing fluorescent proteins. Nat Methods. 2005; 2:905–909. [PubMed: 16299475]

Shcherbakova DM, Hink MA, Joosen L, Gadella TW, Verkhusha VV. An orange fluorescent protein with a large Stokes shift for single-excitation multicolor FCCS and FRET imaging. J Am Chem Soc. 2012; 134:7913–7923. [PubMed: 22486524]

Shcherbo D, Shemiakina II, Ryabova AV, Luker KE, Schmidt BT, Souslova EA, Gorodnicheva TV, Strukova L, Shidlovskiy KM, Britanova OV, Zaraisky AG, Lukyanov KA, Loschenov VB, Luker GD, Chudakov DM. Near-infrared fluorescent proteins. Nat Methods. 2010; 7:827–829. [PubMed: 20818379]

Shi X, Abbyad P, Shu X, Kallio K, Kanchanawong P, Childs W, Remington SJ, Boxer SG. Ultrafast excited-state dynamics in the green fluorescent protein variant S65T/H148D. 2. Unusual photophysical properties. Biochemistry. 2007; 46:12014–12025. [PubMed: 17918960]

Shimomura O. Discovery of green fluorescent protein. Methods Biochem Anal. 2006; 47:1–13. [PubMed: 16335707]

Shu X, Kallio K, Shi X, Abbyad P, Kanchanawong P, Childs W, Boxer SG, Remington SJ. Ultrafast excited-state dynamics in the green fluorescent protein variant S65T/H148D. 1. Mutagenesis and structural studies. Biochemistry. 2007; 46:12005–12013. [PubMed: 17918959]

Shu X, Shaner NC, Yarbrough CA, Tsien RY, Remington SJ. Novel chromophores and buried charges control color in mFruits. Biochemistry. 2006; 45:9639–9647. [PubMed: 16893165]

Shu X, Wang L, Colip L, Kallio K, Remington SJ. Unique interactions between the chromophore and glutamate 16 lead to far-red emission in a red fluorescent protein. Protein Sci. 2009; 18:460–466. [PubMed: 19165727]

Sniegowski JA, Lappe JW, Patel HN, Huffman HA, Wachter RM. Base catalysis of chromophore formation in Arg96 and Glu222 variants of green fluorescent protein. J Biol Chem. 2005; 280:26248–26255. [PubMed: 15888441]

Staiano M, D'Auria S, Varriale A, Rossi M, Marabotti A, Fini C, Stepanenko OV, Kuznetsova IM, Turoverov KK. Stability and dynamics of the porcine odorant-binding protein. Biochemistry. 2007; 46:11120–11127. [PubMed: 17845011]

Stepanenko, OV.; Fonin, AV.; Stepanenko, OV.; Kuznetsova, IM.; Turoverov, KK. Ligand-binding proteins: structure, stability and practical application. In: Faraggi, E., editor. Protein Structure. InTech; Rijeka: 2012a. p. 265-290.

Stepanenko OV, Stepanenko OV, Kuznetsova IM, Shcherbakova DM, Verkhusha VV, Turoverov KK. Distinct effects of guanidine thiocyanate on the structure of superfolder GFP. PLoS One. 2012b; 7(11):e48809. [PubMed: 23144981]

Stepanenko OV, Verkhusha VV, Kazakov VI, Shavlovsky MM, Kuznetsova IM, Uversky VN, Turoverov KK. Comparative studies on the structure and stability of fluorescent proteins EGFP, zFP506, mRFP1, "dimer2", and DsRed1. Biochemistry. 2004; 43:14913–14923. [PubMed: 15554698]

Stepanenko OV, Verkhusha VV, Kuznetsova IM, Uversky VN, Turoverov KK. Fluorescent proteins as biomarkers and biosensors: throwing color lights on molecular and cellular processes. Curr Protein Pept Sci. 2008; 9:338–369. [PubMed: 18691124]

Stiel AC, Andresen M, Bock H, Hilbert M, Schilde J, Schonle A, Eggeling C, Egner A, Hell SW, Jakobs S. Generation of monomeric reversibly switchable red fluorescent proteins for far-field fluorescence nanoscopy. Biophys J. 2008; 95:2989–2997. [PubMed: 18658221]

Subach FV, Malashkevich VN, Zencheck WD, Xiao H, Filonov GS, Almo SC, Verkhusha VV. Photoactivation mechanism of PAmCherry based on crystal structures of the protein in the dark and fluorescent states. Proc Natl Acad Sci USA. 2009a; 106:21097–21102. [PubMed: 19934036]

Subach FV, Patterson GH, Manley S, Gillette JM, Lippincott-Schwartz J, Verkhusha VV. Photoactivatable mCherry for high-resolution two-color fluorescence microscopy. Nat Methods. 2009b; 6:153–159. [PubMed: 19169259]

Subach FV, Subach OM, Gundorov IS, Morozova KS, Piatkevich KD, Cuervo AM, Verkhusha VV. Monomeric fluorescent timers that change color from blue to red report on cellular trafficking. Nat Chem Biol. 2009c; 5:118–126. [PubMed: 19136976]

Subach FV, Patterson GH, Renz M, Lippincott-Schwartz J, Verkhusha VV. Bright monomeric photoactivatable red fluorescent protein for two-color super-resolution sptPALM of live cells. J Am Chem Soc. 2010a; 132:6481–6491. [PubMed: 20394363]

Subach FV, Zhang L, Gadella TW, Gurskaya NG, Lukyanov KA, Verkhusha VV. Red fluorescent protein with reversibly photoswitchable absorbance for photochromic FRET. Chem Biol. 2010b; 17:745–755. [PubMed: 20659687]

Subach OM, Malashkevich VN, Zencheck WD, Morozova KS, Piatkevich KD, Almo SC, Verkhusha VV. Structural characterization of acylimine-containing blue and red chromophores in mTagBFP and TagRFP fluorescent proteins. Chem Biol. 2010c; 17:333–341. [PubMed: 20416505]

Subach FV, Verkhusha VV. Chromophore transformations in red fluorescent proteins. Chem Rev. 2012; 112:4308–4327. [PubMed: 22559232]

Subach OM, Cranfill PJ, Davidson MW, Verkhusha VV. An enhanced monomeric blue fluorescent protein with the high chemical stability of the chromophore. PLoS One. 2011a; 6:e28674. [PubMed: 22174863]

Subach OM, Patterson GH, Ting LM, Wang Y, Condeelis JS, Verkhusha VV. A photoswitchable orange-to-far-red fluorescent protein, PSmOrange. Nat Methods. 2011b; 8:771–777. [PubMed: 21804536]

Subach OM, Entenberg D, Condeelis JS, Verkhusha VV. A FRET-facilitated photoswitching using an orange fluorescent protein with the fast photoconversion kinetics. J Am Chem Soc. 2012; 134:14789–14799. [PubMed: 22900938]

Subach OM, Gundorov IS, Yoshimura M, Subach FV, Zhang J, Gruenwald D, Souslova EA, Chudakov DM, Verkhusha VV. Conversion of red fluorescent protein into a bright blue probe. Chem Biol. 2008; 15:1116–1124. [PubMed: 18940671]

Sulatskaya AI, Kuznetsova IM, Turoverov KK. Interaction of thioflavin T with amyloid fibrils: stoichiometry and affinity of dye binding, absorption spectra of bound dye. J Phys Chem B. 2011; 115:11519–11524. [PubMed: 21863870]

Sulatskaya AI, Povarova OI, Kuznetsova IM, Uversky VN, Turoverov KK. Binding stoichiometry and affinity of fluorescent dyes to proteins in different structural states. Methods Mol Biol. 2012; 895:441–460. [PubMed: 22760333]

Teerawanichpan P, Hoffman T, Ashe P, Datla R, Selvaraj G. Investigations of combinations of mutations in the jellyfish green fluorescent protein (GFP) that afford brighter fluorescence, and use of a version (VisGreen) in plant, bacterial, and animal cells. Biochim Biophys Acta. 2007; 1770:1360–1368. [PubMed: 17658219]

Tinnefeld P, Sauer M. Branching out of single-molecule fluorescence spectroscopy: challenges for chemistry and influence on biology. Angew Chem Int Ed Engl. 2005; 44:2642–2671. [PubMed: 15849689]

Tsien RY. The green fluorescent protein. Annu Rev Biochem. 1998; 67:509–544. [PubMed: 9759496]

Tsutsui H, Karasawa S, Shimizu H, Nukina N, Miyawaki A. Semi-rational engineering of a coral fluorescent protein into an efficient highlighter. EMBO Rep. 2005; 6:233–238. [PubMed: 15731765]

Turoverov KK, Kuznetsova IM, Uversky VN. The protein kingdom extended: ordered and intrinsically disordered proteins, their folding, supramolecular complex formation, and aggregation. Prog Biophys Mol Biol. 2010; 102:73–84. [PubMed: 20097220]

Ugrinov KG, Clark PL. Cotranslational folding increases GFP folding yield. Biophys J. 2010; 98:1312–1320. [PubMed: 20371331]

Uversky VN, Dunker AK. Understanding protein non-folding. Biochim Biophys Acta. 2010; 1804:1231–1264. [PubMed: 20117254]

Uversky VN, Cooper ME, Bower KS, Li J, Fink AL. Accelerated alpha-synuclein fibrillation in crowded milieu. FEBS Lett. 2002; 515:99–103. [PubMed: 11943202]

Uversky VN, Ptitsyn OB. Further evidence on the equilibrium “pre-molten globule state”: four-state guanidinium chloride-induced unfolding of carbonic anhydrase B at low temperature. J Mol Biol. 1996; 255:215–228. [PubMed: 8568868]

van den Berg B, Wain R, Dobson CM, Ellis RJ. Macromolecular crowding perturbs protein refolding kinetics: implications for folding inside the cell. EMBO J. 2000; 19:3870–3875. [PubMed: 10921869]

van den Berg S, Lofdahl PA, Hard T, Berglund H. Improved solubility of TEV protease by directed evolution. J Biotechnol. 2006; 121:291–298. [PubMed: 16150509]

van Thor JJ, Georgiev GY, Towrie M, Sage JT. Ultrafast and low barrier motions in the photoreactions of the green fluorescent protein. J Biol Chem. 2005; 280:33652–33659. [PubMed: 16033764]

Verkhusha VV, Kuznetsova IM, Stepanenko OV, Zaraisky AG, Shavlovsky MM, Turoverov KK, Uversky VN. High stability of Discosoma DsRed as compared to *Aequorea* EGFP. Biochemistry. 2003; 42:7879–7884. [PubMed: 12834339]

Verkhusha VV, Lukyanov KA. The molecular properties and applications of *Anthozoa* fluorescent proteins and chromoproteins. Nat Biotechnol. 2004; 22:289–296. [PubMed: 14990950]

Vrzheshch PV, Akovbian NA, Varfolomeyev SD, Verkhusha VV. Denaturation and partial renaturation of a tightly tetramerized DsRed protein under mildly acidic conditions. FEBS Lett. 2000; 487:203–208. [PubMed: 11150510]

Wachter RM, Elsliger MA, Kallio K, Hanson GT, Remington SJ. Structural basis of spectral shifts in the yellow-emission variants of green fluorescent protein. Structure. 1998; 6:1267–1277. [PubMed: 9782051]

Wachter RM, Remington SJ. Sensitivity of the yellow variant of green fluorescent protein to halides and nitrate. Curr Biol. 1999; 9:R628–R629. [PubMed: 10508593]

Waldo GS, Standish BM, Berendzen J, Terwilliger TC. Rapid protein-folding assay using green fluorescent protein. Nat Biotechnol. 1999; 17:691–695. [PubMed: 10404163]

Ward WW, Bokman SH. Reversible denaturation of *Aequorea* green-fluorescent protein: physical separation and characterization of the renatured protein. Biochemistry. 1982; 21:4535–4540. [PubMed: 6128025]

Ward WW, Prentice HJ, Roth AF, Cody CW, Reeves SC. Spectral perturbations of the *Aequorea* Green-Fluorescent protein. Photochem Photobiol. 1982; 35:803–808.

White A. Effect of pH on fluorescence of tryosine, tryptophan and related compounds. Biochem J. 1959; 71:217–220. [PubMed: 13628557]

Wiedenmann J, Gayda S, Adam V, Oswald F, Nienhaus K, Bourgeois D, Nienhaus GU. From EosFP to mIrisFP: structure-based development of advanced photo-activatable marker proteins of the GFP-family. J Biophotonics. 2011; 4:377–390. [PubMed: 21319305]

Wiedenmann J, Ivanchenko S, Oswald F, Schmitt F, Rocker C, Salih A, Spindler KD, Nienhaus GU. EosFP, a fluorescent marker protein with UV-inducible green-to-red fluorescence conversion. Proc Natl Acad Sci USA. 2004; 101:15905–15910. [PubMed: 15505211]

Wiedenmann J, Vallone B, Renzi F, Nienhaus K, Ivanchenko S, Rocker C, Nienhaus GU. Red fluorescent protein eqFP611 and its genetically engineered dimeric variants. J Biomed Opt. 2005; 10:14003. [PubMed: 15847584]

Wilmann PG, Petersen J, Pettikiriarachchi A, Buckle AM, Smith SC, Olsen S, Perugini MA, Devenish RJ, Prescott M, Rossjohn J. The 2.1A crystal structure of the far-red fluorescent protein HcRed: inherent conformational flexibility of the chromophore. J Mol Biol. 2005; 349:223–237. [PubMed: 15876379]

Wood TI, Barondeau DP, Hitomi C, Kassmann CJ, Tainer JA, Getzoff ED. Defining the role of arginine 96 in green fluorescent protein fluorophore biosynthesis. Biochemistry. 2005; 44:16211–16220. [PubMed: 16331981]

Xie JB, Zhou JM. Trigger factor assisted folding of green fluorescent protein. Biochemistry. 2008; 47:348–357. [PubMed: 18067273]

Yampolsky IV, Remington SJ, Martynov VI, Potapov VK, Lukyanov S, Lukyanov KA. Synthesis and properties of the chromophore of the asFP595 chromoprotein from *Anemonia sulcata*. Biochemistry. 2005; 44:5788–5793. [PubMed: 15823037]

Yang F, Moss LG, Phillips GN Jr. The molecular structure of green fluorescent protein. Nat Biotechnol. 1996; 14:1246–1251. [PubMed: 9631087]

Yanushevich YG, Staroverov DB, Savitsky AP, Fradkov AF, Gurskaya NG, Bulina ME, Lukyanov KA, Lukyanov SA. A strategy for the generation of non-aggregating mutants of *Anthozoa* fluorescent proteins. FEBS Lett. 2002; 511:11–14. [PubMed: 11821040]

Yarbrough D, Wachter RM, Kallio K, Matz MV, Remington SJ. Refined crystal structure of DsRed, a red fluorescent protein from coral, at 2.0-A resolution. Proc Natl Acad Sci USA. 2001; 98:462–467. [PubMed: 11209050]

Yokoyama S. Protein expression systems for structural genomics and proteomics. Curr Opin Chem Biol. 2003; 7:39–43. [PubMed: 12547425]

Zacharias DA, Violin JD, Newton AC, Tsien RY. Partitioning of lipid-modified monomeric GFPs into membrane microdomains of live cells. Science. 2002; 296:913–916. [PubMed: 11988576]

Zimmerman SB, Minton AP. Macromolecular crowding: biochemical, biophysical, and physiological consequences. Annu Rev Biophys Biomol Struct. 1993; 22:27–65. [PubMed: 7688609]

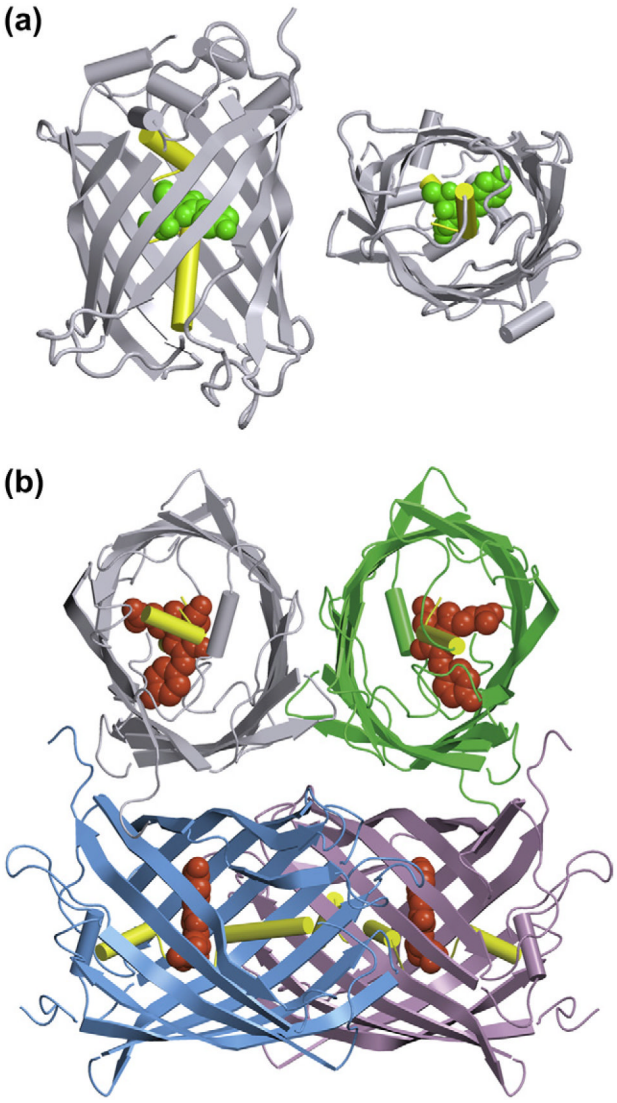

Figure 4.1. Three-dimensional structure of sfGFP (PDB code 2B3P, Pedelacq et al., 2006) in two projections

(*a*) and of DsRed1 from *Discosoma* sp. (PDB code 1G7K, Yarbrough et al., 2001) (*b*). The chromophores of sfGFP and DsRed1 are shown as *green* and *red space-filling unions*, respectively. A central α-helix bearing the chromophore is shown in *yellow*. Monomers of DsRed1 are displayed in different colors. The drawing was generated by the graphic programs VMD (Humphrey et al., 1996) and Raster3D (Merritt and Bacon, 1997). (For interpretation of the references to color in this figure legend, the reader is referred to the online version of this book).

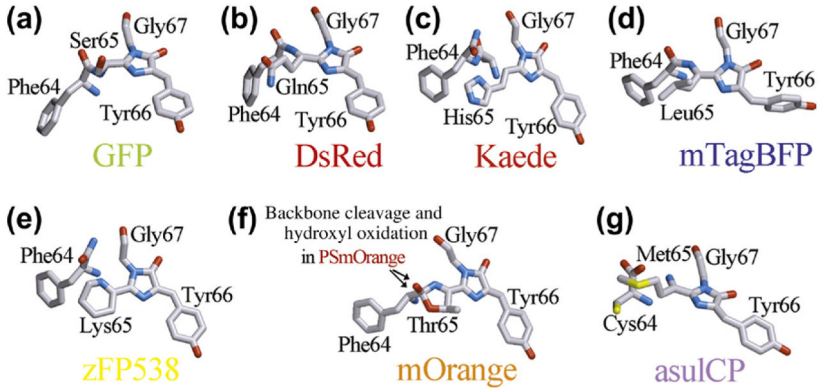

Figure 4.2. A variety of chromophore structures in FPs

a – green chromophore of GFP (PDB code 1W7S; van Thor et al., 2005); *b* and *c* – red chromophores of DsRed (PDB code 1G7K; Yarbrough et al., 2001) and Kaede (PDB code 2GW4; Hayashi et al., 2007); *d* – blue chromophore of mTagBFP (PDB code 3M24; Subach et al., 2010c); *e*–*g* – derivatives of the DsRed-like red chromophore of zFP538 (PDB code 1XAE; Remington et al., 2005), mOrange (PDB code 2H5O; Shu et al., 2006), PSmOrange and asulCP (PDB code 2A50; Andresen et al., 2005). Carbon, nitrogen, oxygen and sulfur are colored in gray, blue, red and yellow, respectively. The drawing was generated based on the Protein Data Bank (Dutta et al., 2009) by the graphic programs VMD (Humphrey et al., 1996) and Raster3D (Merritt and Bacon, 1997). (For interpretation of the references to color in this figure legend, the reader is referred to the online version of this book).

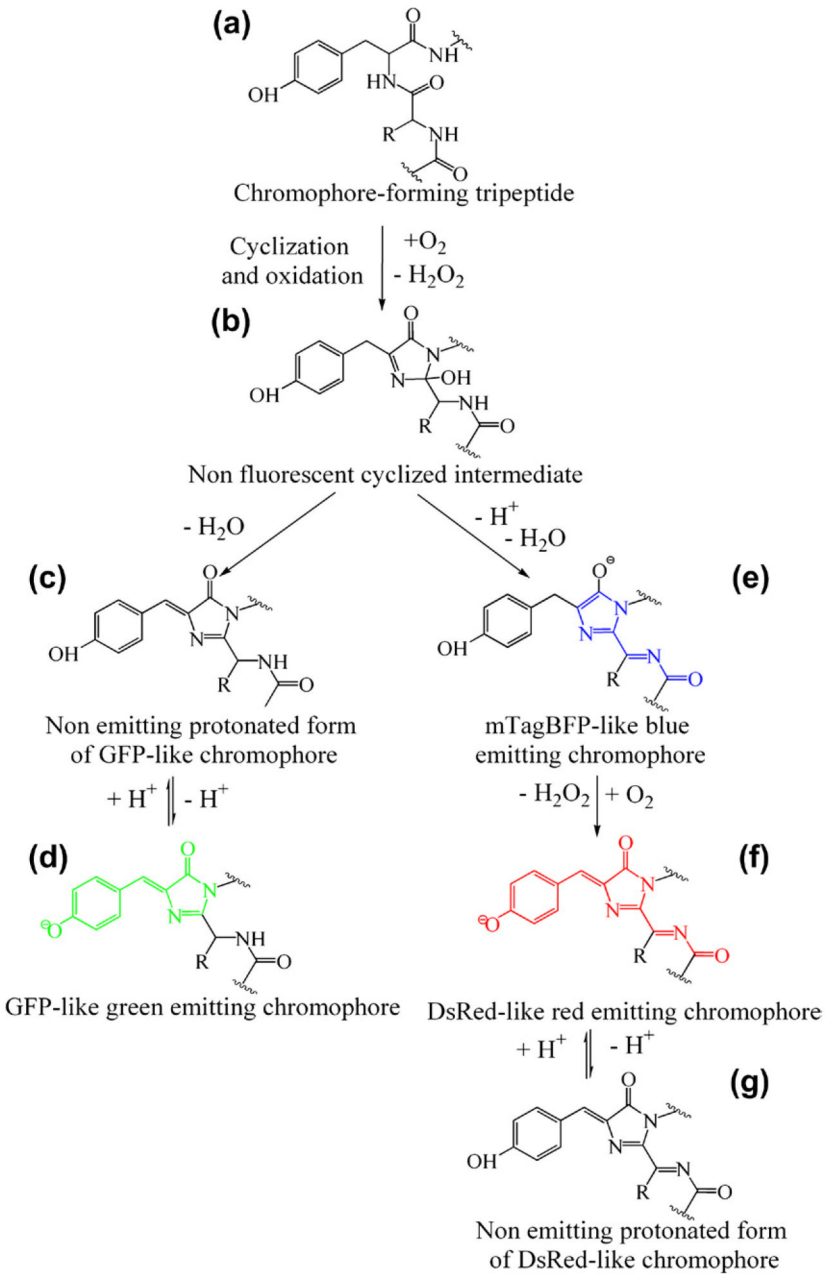

Figure 4.3.

General scheme of the autocatalytic synthesis of blue, green and red chromophores. (For interpretation of the references to color in this figure legend, the reader is referred to the online version of this book).

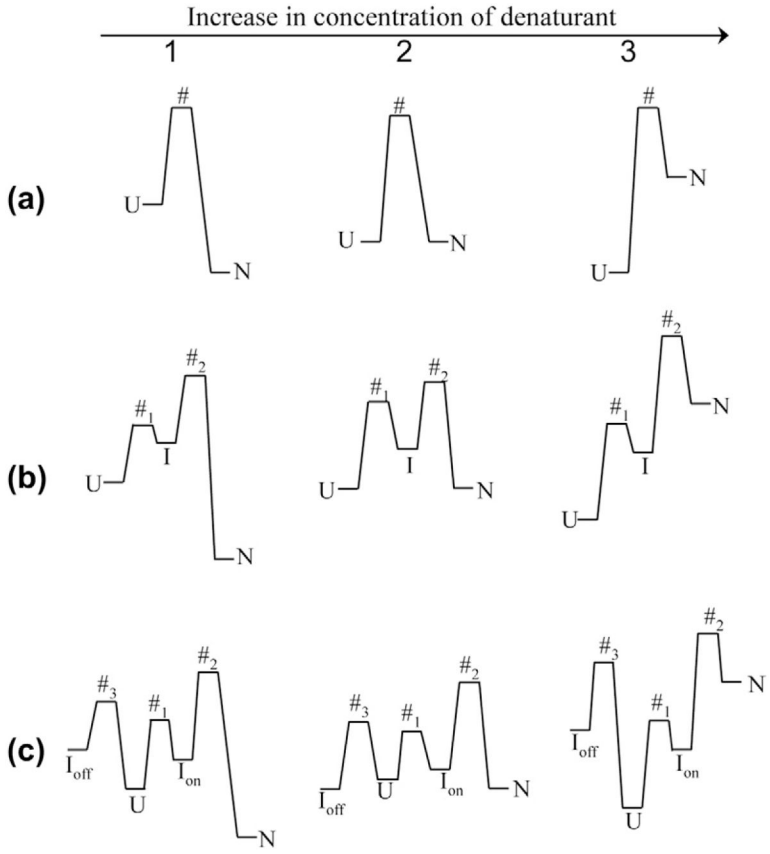

Figure 4.4. Change in the energy landscape induced by a denaturant

1 and 3 are the concentrations of denaturant at which the protein is in its native and unfolded states. U, N, I and # are the unfolded, native, intermediate and transition states of protein. I_{on} and I_{off} are the on- and off-pathway intermediate states. *a*. Two-state unfolding–refolding model of protein. *b*. Three-state unfolding–refolding model of protein. *c*. Protein unfolding–refolding via on- and off-pathway intermediates.

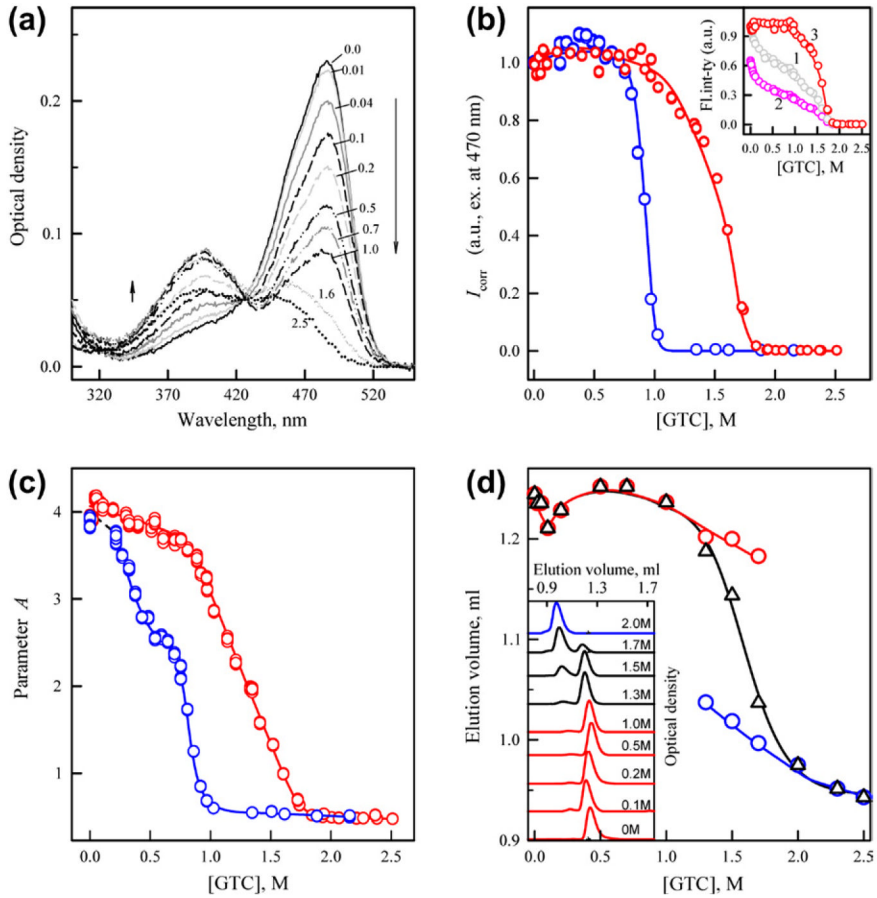

Figure 4.5. sfGFP unfolding–refolding induced by GTC (Stepanenko et al., 2012b)

a, change in the absorption spectrum; *b*, chromophore fluorescence intensity corrected to the change of the chromophore absorption spectra and *c*, parameter $A = I_{320}/I_{356}$ of tryptophan fluorescence on GTC concentration. Inset to panel *b*: experimentally recorded chromophore fluorescence intensity (curve 1, gray), corrected to a total density of solution as follows I/W, where $W = (1 - 10^{-D_\Sigma})/D_\Sigma$ (see Kuznetsova et al., 2012; Sulatskaya et al., 2011; Sulatskaya et al., 2012) (curve 2, pink), and corrected to the change of the chromophore absorption spectra (panel a) with the GTC concentration (curve 3, red). *d*, changes of the position of elution peaks of compact and denatured molecules (red and blue circles, respectively) and the change of the averaged elution volume of sfGFP (black triangles). Inset: Changes of the elution profile of sfGFP at increasing denaturant concentrations. The values of the curves specify applied denaturant concentration. (For interpretation of the references to color in this figure legend, the reader is referred to the online version of this book).

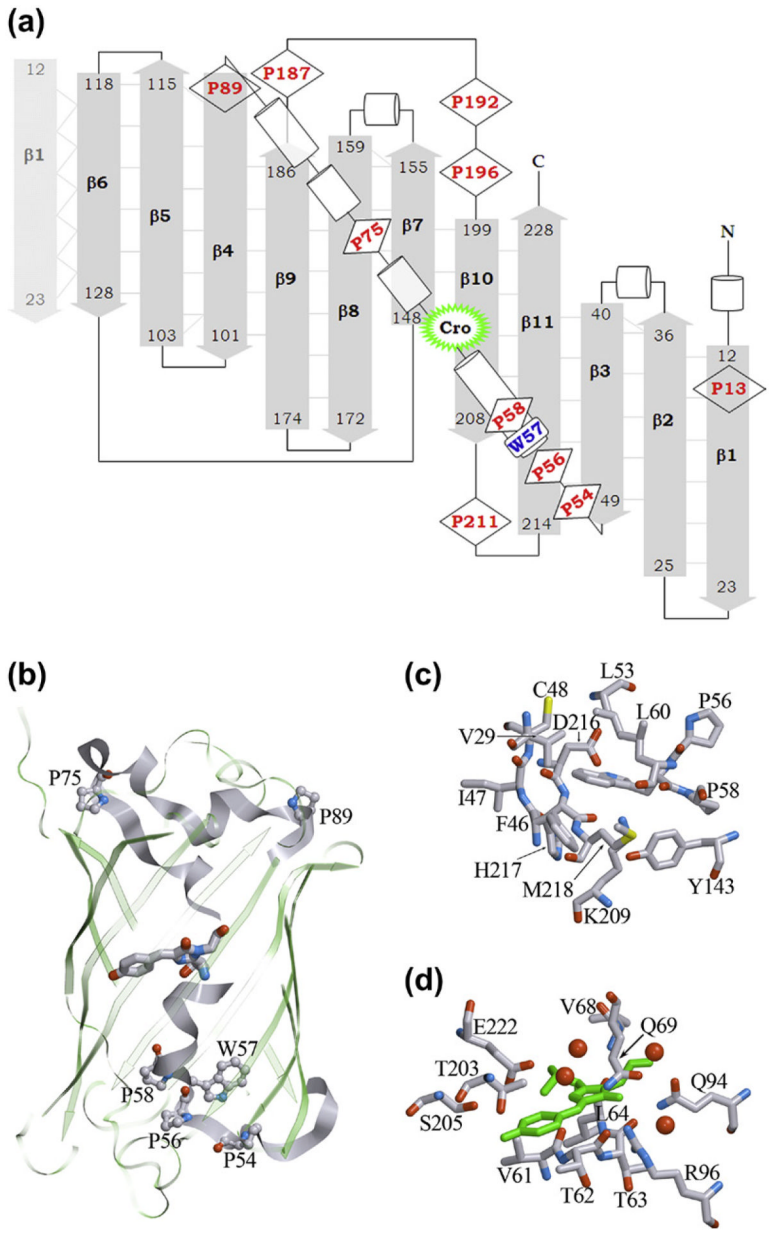

Figure 4.6. Structure of sfGFP

Localization and microenvironment of the chromophore and the tryptophan residue. *a*, diagram illustrating the formation of the β-barrel of 11 β-strands and the internal helix. The localization of the chromophore (Cro), tryptophan residue W57 and proline residues, including Pro89, the only proline that has the *cis*-conformation is shown; *b*, localization of α-helix in β-barrel. The β-barrel strand in the foreground is made transparent. The proline residues that are part of the α-helix, and Pro89, which is localized between the α-helix and the fourth β-strand, are shown; *c*, the microenvironment of W57; *d*, the chromophore microenvironment. (For interpretation of the references to color in this figure legend, the reader is referred to the online version of this book).

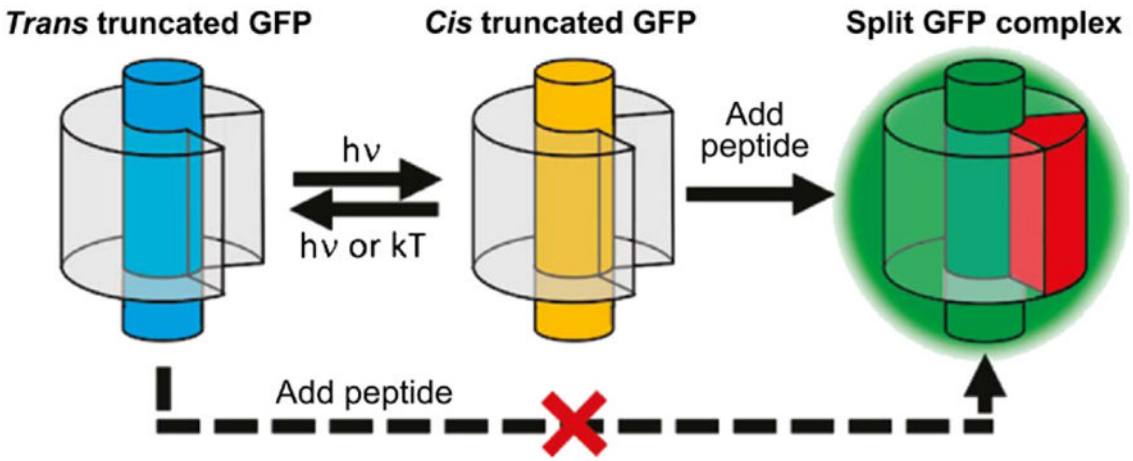

Figure 4.7. Reassembling sfGFP 1–10 and a synthetic 11-th β-strand

The truncated sfGFP 1–10 after refolding in native conditions does not reassemble with a synthetic peptide corresponding to 11-th β-strand, but it reassembles after light activation. The reassembled structure is identical to the native protein *(Reprinted with permission from (* Kent and Boxer, 2011*) Copyright 2011 American Chemical Society).* (For color version of this figure, the reader is referred to the online version of this book).

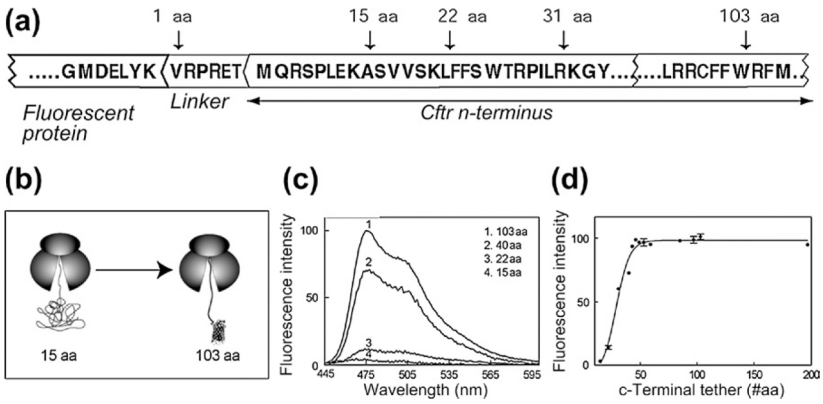

Figure 4.8. Co-translational folding of FPs

a, a diagram of the FP fusion protein showing the C-terminal of the FP, the 6-amino acids linker and the N-terminal residues of the cystic fibrosis transmembrane conductance regulator (CFTR). The truncation sites are indicated by arrows. *b*, a diagram showing co-translational folding of the FP with different C-terminal tether lengths. *c*, the emission fluorescence spectra of the FP with different C-terminal tether lengths. *d*, the dependence of the intensity fluorescence of FP with different C-terminal tether lengths. *(Adapted from* figure 1 and figure 2 originally published in Journal of Biological Chemistry (Kelkar et al., 2012) Copyright 2012 the American Society for Biochemistry and Molecular Biology).