Bring to Structure to Green Fluorescent Protein

BY YUEJIAN MO

Southern University of Science and Technology
June 24, 2018

Introduction

In the1960 and 1970s, *Aequorea vitoria* green fluorescent protiein(avGFP), along with the separate luminescent protein aequorin was first purified from *Aequorea victoria* and its properties studied by Osamu Shimomura.[7] Then the lab of Martin Chalfie expressed the modfied coding sequence of avGFP in *E.coli* and *C. elegans*, publishing in *Science* in 1994. [2]. Due to the potential for widespread usage, many different mutants of GFP have been engineered. The first major improvement was a single point mutation(S65T) reported in 1995 in Nature by Roger Tsien[5]. The first reported crystal stcucture of GFP, is mutant S65T by the Remington group in 1996[6]. These crystal structure of GFP can suggeste chromophore formation and neighboring redidue interactions, which points out possible functional residue for more modification. Nowaday, wide variety of GFP derivatives are developed and used in laboratories.

Fluorescent protiens paly lots of important role in wide biological field, such as genetic marker, indicator, even contribution to super reslution microscopy. Fluorescence is most exciting and unique characteristic of GFP. Different GFP varitants show different spectra, dymacis, envirnoment resistant and sensablity and etc. Based on these feactures, suitable GFPs vaiants are used in special experiments. In other side, new GFP vaiants or other fluorescent protein are devoloped for new experiment

requirement.

So, in this article, by comparing existen GFP vaiants and other fluorescent protein, we try to figure out how the structure function on different features, escipcially in spectra.

The Structure and Formation of wtGFP

GFP family consist of ~220-240 amino acid residues(25kDa), which fold in to a barrel formed by 11-βsheets that accommodates an internal distorsted helix. The chromophore is attached to the α-helix and is buried almost perfectly in the center of the cylinder(Figure.1). The barrel of green fluorescent proteins is stabilized by mulitple noncovalent interactions that ensure its extremely high stability to thermal or chemical denaturation as well as resistance to proteolysis.

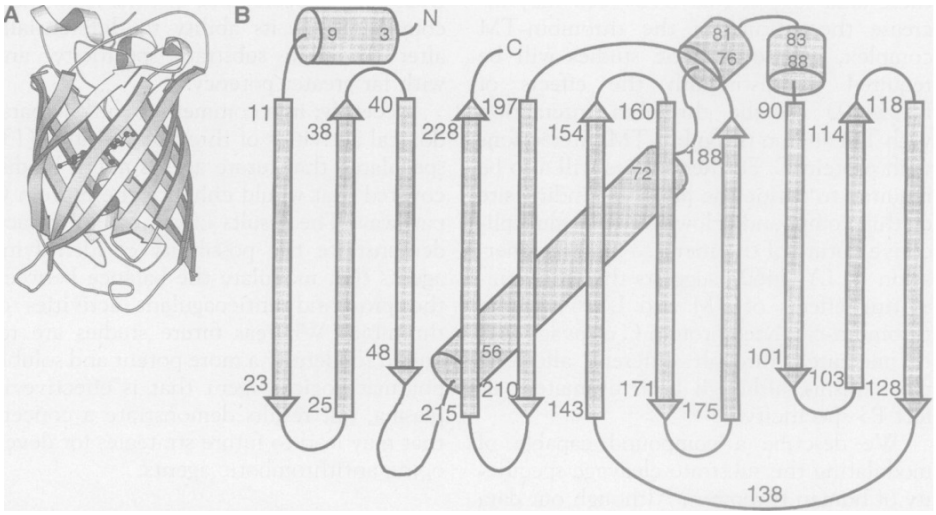

Figure 1. (A)Schematic drawing of the backbone of GFP. (B) Schematic drawing of the overall fold of GFP.[6]

The chromophore is most interesting part in green fluorescent protein, which is result after cyclizatino and oxygenation by catalyticed by β-barrel in posttranslational modification. Commonly, biosynthesis of pigments occurs through multiple sequential reactions catalyzed by severl specific enzymes and involving low-molecular substrates and cofactors. However, the barrel of GFP play a role as a unique “enzyme” that modifies its own internal amino acids without any external cofactors and substrates except molecular oxygen. The chromophore is a *p*-hydroxybenzylideneimida-

zolinone formed from residues 65-67(numbering in accordance with *Aequoresa victoria* GFP), which are Ser-Tyr-Gly in the native protein. Figure 3 show a accepted mechanism for chromophore formation. First, GFP folds into a nearly native conformation, then the imidazolinone is formed by nucleophilic attack of the amide of Gly67 on the carbonyl of residue 65, followed by dehydration. Finally, molecular oxygen dehydrogenates the α-β bond of residue 66 to put its aromatic group into conjugation with the imidazolinone. However, the exact sequence of molecular events during chromophore maturation is still a matter of active debate.

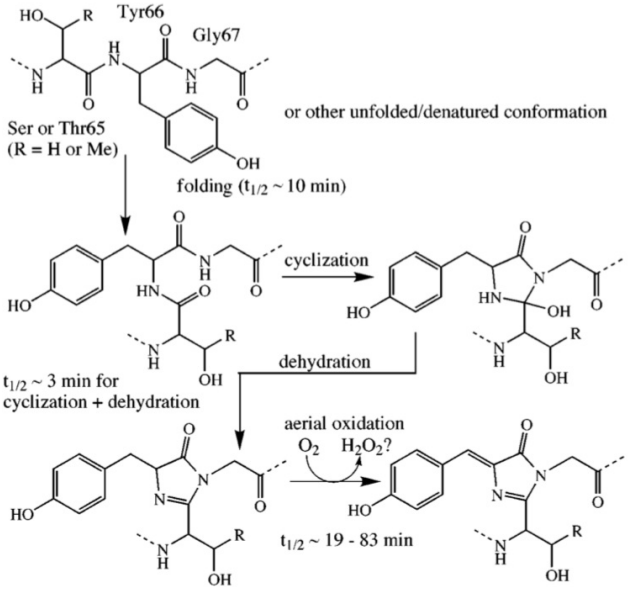

Figure 2. Mechanism proposed by Cubbitt et al for the intramolecular biosynthesis of the GFP chromophore, with rate constants estimated for the Ser65 → Thr mutant by Reid & Flynn and Heim et al.[8]

The resulting chromophore is located in the very center of the β-barrel and therefore is well protected from contact with the solvent by the surrounding protien shell.(Figure.3) Besides the three residues that form the chromophore, residues such as Gln94, Arg96, His148, Thr203, and Glu222 all act as stabilizers. The residues of Gln94, Arg96, and His148 are able to stabilize by delocalizing the chromophore charge.

General Relation of Structure to Spectra

Most exciting and unique characteristic is the fluorescence of green fluorescence pro-

tein. Genraly, when exciting light light on chromophore, electrons in tripletground state on chromophore abosrb this energy and transite to higher energy level. After a short time, excited electrons down to low energy level while extra energy is relased in lower frequency light called emission light. Different chromophore and interaction with surrouding amino acids function on different exciting and emssion spectra. Here we try to understand what determine the spectra by the difference analysis from avGFP to other GFP variants.

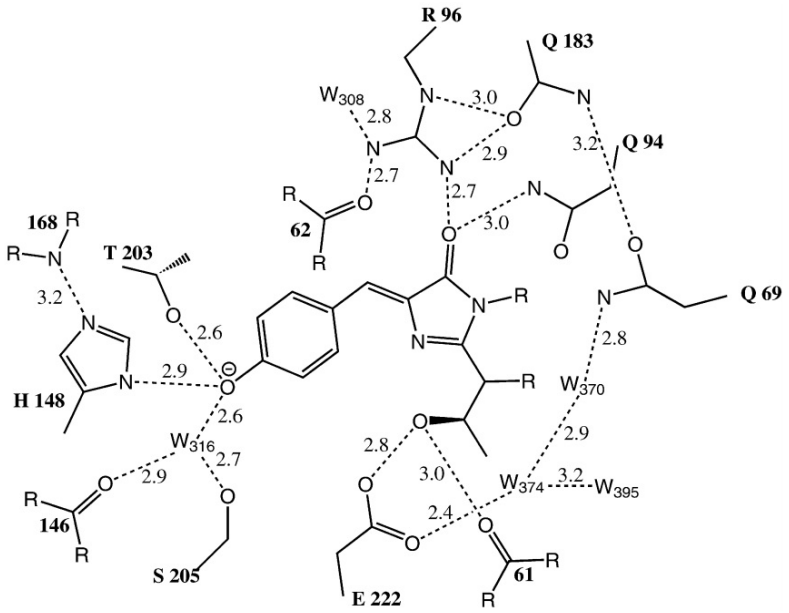

Figure 3. Amino acid side chains, main chian carbonyls and amides, and solvent water in the immediate vicinity of the chromophore of S65T GFP. Figure from SJ Remington, University of Oregon.[6]

The hydrogen-bonding network and electron-stacking interactions with these sidechains influence the color, intensity and photostability of GFP and its numerous derivatives.[3] These reactions result in a two-ring structure representing a conjugated π-system that is large, polarized, and planar enough to absorb and emit light within the visible range.(Figure.4) Then, the tightly packed nature of the barrel excludes solvent molecules, protecting the chromophore-fluorescence from quenching by water dipoles, paramagenetic oxygen molecules, or *cis-trans* isomerization. Denatured GFPs or small proteolytic fragments carrying the chromophore are essentially totally nonfluorescent.[8]

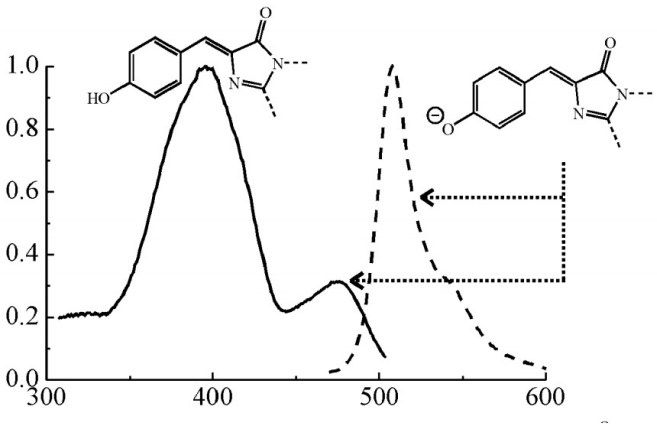

Figure 4. wild type avGFP fluorescene excitation and emission spectra(solid and dashed line, respectively)[8]

The wild type avGFP has several significant deficiencies, including dual peaked excitation spectra(Figure.4), pH sensitivity, chloride sensitivity, poor fluorescen qunatum yield, poor photostability and poor folding at 37°C. The major improvement was a single piont mutation(S65T) by Roger Tsien in 1995. Roger Tsien and others mutated 65Ser to Ala, Leu, Cys or Thr, they found all four mutants showed single excitation peaks, located at 470nm-490 nm, whose amplitudes were four-to sixfold than that of wild type for equal numbers of molecules.[5]

Then a point mutant S65T, and F64L improving 37°C folding efficiency, was discovery in 1995 by the laboratories of Thasturp and Falkow.[4] It called enhanced GFP(EGFP), which is most widely used today. Because EGFP allowed to fold in manmalian cells. The additional mutations F64L may be necessary for efficient folding or chromophore formation; alternately they may have a more direct effect on absorption efficiency.[4]

In wild-type 395nm exicitation peak due to the neutral phenonl, and the 470- to 475-nm peak due to the anion phenonl on Try66.(Figure. 4) Ser65 mutation drecrease ratio of the neutral phenol on Tyr66, which enhance the excitation in amplitude of 489-490nm. The probable mechanism by which replacement of Ser65 promotes Tyr66 ionization is that only Ser65 can donate a hydrogen bond to the buried side chain of Glu222 to allow ionization of that carboxylate, which is within 3.7 A° of the chromophore. Gly, Ala, and Leu cannot donate hydrogen bonds, and Thr and Cys are too large to adopt the correct conformation in the crowded interior of the protein. Such residues at position 65 force the carboxyl of Glu222 to remain neutral. The other polar groups solvating the chromophore are then sufficient to promote its ionization to an anion, whereas if Glu222 is an anion, electrostatic repulsion forbids the chromophore from becoming an anion as well. This hypothesis

explains why mutation of Glu222 to Gly gives the same spectral shape and wavelengths as Ser65 mutations.[8]

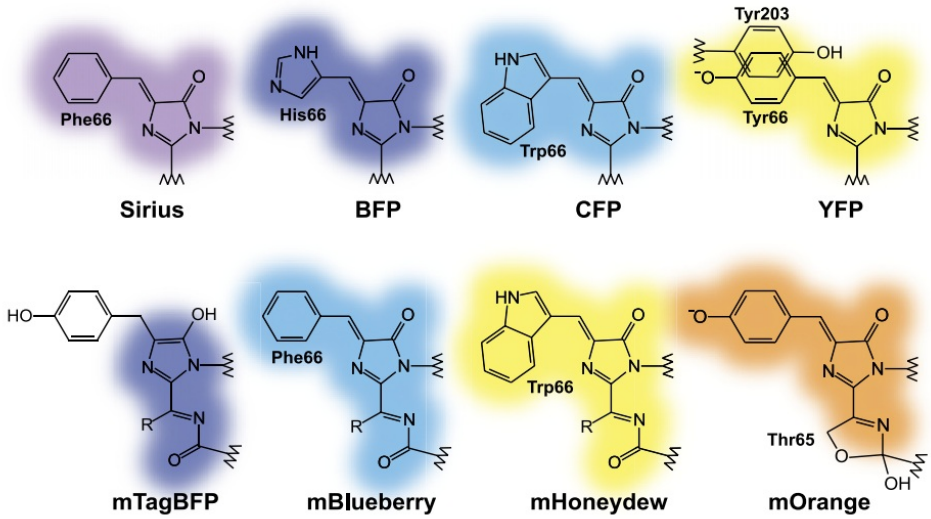

Figure 5. Chemical structures of chromophores obtained in artificially mutated FPs. Top row: blue-shifted chromophores generated by substitution of GFP chromophore's Tyr66 by armatic residus; also, a stacking of Tyr66 with Tyr203 utilized in YFP variants is shown. Blue-shifted modifications of DsRed chromophore obtained by stabilization acylimine-based chromophore, introduction of Phe or Trp residues instead of Tyr66, or by mutation Gln65Thr resulting in cyclization of Thr65 side chain analogous to that of Cys65 in KO.[3]

Excepted for GFP, researchers also devoloped lots of color fluorescent proteins later, such as blue fluorescent protein and yellow fluorescent protein.

Early attempts to mutate avGFP revealed that Tyr66 of the chromophore can be substituted for aromatic side chains(Phe, His, or Trp), resulting in dramatically blue-shifted FPs(Figure.5) Trp-based chromophores give off CFP with an excitation/emission maxima at ~433/475nm. His-based chromophors are characteristic of BFP, which blue-shifited exicitation/emission peaks at 383/448nm. Finally, Phe66 provides the shortest wavelength fluorescence with exicitation peak at 355nm and emission peak at 424nm. It is not clear why these modification don't occuar in natual protein. However, Thr203Tyr from avGFP drived fluorescence red-shift, called yellow FP. A π-π stack between Tyr203 and chromophore's Tyr66, results in significantly red-shifted spectra(515/528nm). In fact, all four aromatic residues at that position 203(His, Trp, Phe and Tyr) increase the exicitation and emission wavelengths by up to 20nm.[8] Additional polarizability around the chromophore and π-π interaction would reduce the excited state energy.

More artificially mutated FPs can be classified in Table1. Above results suggeste that more strong conjugated π-system allow longer exicitation and emission light.

Mutation[a]	Common name	λ_{exc} (ϵ)[b]	λ_{em} (QY)[c]	Rel. fl.[d] @ 37°C	References[e]
Class 1, wild-type					
None or Q80R	Wild type	395–397 (25–30) 470–475 (9.5–14)	504 (0.79)	6	43, 45
F99S, M153T, V163A	Cycle 3	397 (30) 475 (6.5–8.5)	506 (0.79)	100	43, 45
Class 2, phenolate anion					
S65T		489 (52–58)	509–511 (0.64)	12	43–45
F64L, S65T	EGFP	488 (55–57)	507–509 (0.60)	20	43–45
F64L, S65T, V163A		488 (42)	511 (0.58)	54	44
S65T, S72A, N149K, M153T, I167T	Emerald	487 (57.5)	509 (0.68)	100	44
Class 3, neutral phenol					
S202F, T203I	H9	399 (20)	511 (0.60)	13	44
T203I, S72A, Y145F	H9–40	399 (29)	511 (0.64)	100	44
Class 4, phenolate anion with stacked π-electron system (yellow fluorescent proteins)					
S65G, S72A, T203F		512 (65.5)	522 (0.70)	6	44
S65G, S72A, T203H		508 (48.5)	518 (0.78)	12	44
S65G, V68L, Q69K S72A, T203Y	10C Q69K	516 (62)	529 (0.71)	50	44
S65G, V68L, S72A, T203Y	10C	514 (83.4)	527 (0.61)	58	44
S65G, S72A, K79R, T203Y	Topaz	514 (94.5)	527 (0.60)	100	44
Class 5, indole in chromophore (cyan fluorescent proteins)					
Y66W		436	485	—	21
Y66W, N146I, M153T, V163A	W7	434 (23.9) 452	476 (0.42) 505	61	44
F64L, S65T, Y66W, N146I, M153T, V163A	W1B or ECFP	434 (32.5) 452	476 (0.4) 505	80	44
S65A, Y66W, S72A, N146I, M153T, V163A	W1C	435 (21.2)	495 (0.39)	100	44
Class 6, imidazole in chromophore (blue fluorescent proteins)					
Y66H	BFP	384 (21)	448 (0.24)	18	44
Y66H, Y145F	P4–3	382 (22.3)	446 (0.3)	52	44
F64L, Y66H, Y145F	EBFP	380–383 (26.3–31)	440–447 (0.17–0.26)	100	43, 44
Class 7, phenyl in chromophore					
Y66F		360	442	—	22

Table 1. Spectral characteristics of the major classes of GFPs.[8]

During developing practice-friendly, mulit-purpose and colorful fluorescent protein, research also find out following facts.

Arg96 is the most important stabilizing residue due to the fact that it prompts the necessary structural realignments that are necessary from the HBI ring to occur. Any mutation to the Arg96 residue would result in a decrease in the development rate of the chromophore because proper electrostatic and steric interactions would be lost.[3]

Then, Arg96 puts a positively charged guanidinium quite close to the carbonyl group of the imidazolinone. This cation would electrostatically stabilize incread electron-density on the carbonyl oxygen in the chromophore's excited state. This electrostatic attration would explain much of the red shift of intact protein relative to denatured protein. Indeed, mutation of Arg96 to Cys in S65T blue-shifts the exci-

tation maximum from 489 to 472nm and the emission maximum for 511 to 503 nm, supporting a major role for Arg96 in lowering the energy of the excited state.[8]

Conclusions

From the strucuture and formation of GFP and its variant, we take a look how hydrogen-bonding network and electron-stacking interactions with these sidechains influence the spetra, intensity of GFP. However, it is pity that we are difficult to caluculate the spectra from structure. It will be amazing and inspired if we can do it in future work.

Wikipideia update: https://zh.wikipedia.org/w/index.php?title=%E7%BB%BF%E8%89%B2%E
tory ID: JiangXL

Bibliography

[1] Katjuša Brejc, Titia K. Sixma, Paul A. Kitts, Steven R. Kain, Roger Y. Tsien, Mats Ormö, and S. James Remington. Structural basis for dual excitation and photoisomerization of the Aequorea victoria green fluorescent protein. *Proceedings of the National Academy of Sciences of the United States of America*, 94(6):2306–2311, mar 1997.

[2] M Chalfie, Y Tu, G Euskirchen, W. Ward, and D. Prasher. Green fluorescent protein as a marker for gene expression. *Science*, 263(5148):802–805, feb 1994.

[3] Dmitriy M. Chudakov, Mikhail V. Matz, Sergey Lukyanov, and Konstantin A. Lukyanov. Fluorescent Proteins and Their Applications in Imaging Living Cells and Tissues. *Physiological Reviews*, 90(3):1103–1163, jul 2010.

[4] B. P. Cormack, R. H. Valdivia, and S. Falkow. FACS-optimized mutants of the green fluorescent protein (GFP). *Gene*, 173(1 Spec No):33–38, 1996.

[5] R. Heim, A. B. Cubitt, and R. Y. Tsien. Improved green fluorescence. *Nature*, 373(6516):663–664, feb 1995.

[6] M. Ormö, A. B. Cubitt, K. Kallio, L. A. Gross, R. Y. Tsien, and S. J. Remington. Crystal structure of the Aequorea victoria green fluorescent protein. *Science (New York, N.Y.)*, 273(5280):1392–1395, sep 1996.

[7] Osamu Shimomura, Frank H. Johnson, and Yo Saiga. Extraction, Purification and Properties of Aequorin, a Bioluminescent Protein from the Luminous Hydromedusan,Aequorea. *Journal of Cellular and Comparative Physiology*, 59(3):223–239, jun 1962.

[8] Roger Y. Tsien. THE GREEN FLUORESCENT PROTEIN. *Annual Review of Biochemistry*, 67(1):509–544, jun 1998.

[9] Jin Zhang, Robert E. Campbell, Alice Y. Ting, and Roger Y. Tsien. Creating new fluorescent probes for cell biology. *Nature Reviews Molecular Cell Biology*, 3(12):906–918, dec 2002.