J|A|C|S
COMMUNICATIONS

Published on Web 02/16/2006

Structural Evidence for an Enolate Intermediate in GFP Fluorophore Biosynthesis

David P. Barondeau, John A. Tainer, and Elizabeth D. Getzoff*

Department of Molecular Biology, The Skaggs Institute for Chemical Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037

Received August 11, 2005; E-mail: edg@scripps.edu

The *Aequorea victoria* green fluorescent protein (GFP) creates a fluorophore from its component amino acids Ser65, Tyr66, and Gly67 by driving backbone cyclization and oxidation post-translational chemistry without any assisting cofactors or enzymes.[1−4] The GFP chromophore is formed in a distorted α-helix buried within an 11-stranded antiparallel β-barrel protein (see Figure 3C) through three primary reaction steps (Figure 1): peptide backbone cyclization initiated by nucleophilic attack of the Gly67 amide nitrogen on the Ser65 carbonyl carbon to create a five-membered imidazolone ring; dehydration of the Ser65 carbonyl group; and oxidation of the Tyr66 Cα−Cβ bond to conjugate the ring systems.[2−5] We previously examined the determinants for backbone cyclization by isolating and characterizing structures of precursor intermediate and mature chromophore states for the R96A and S65G Y66G GFP variants.[6] Our analyses suggested that the protein scaffold promotes backbone cyclization by eliminating inhibitory main-chain hydrogen bonds in the precursor state, aligning the Gly67 amide lone pair with the π*-orbital of the residue 65 carbonyl for bond formation, and then trapping or stabilizing a unfavorable cyclization product through subsequent post-translational chemistry.[6,7] Here we couple chemical reduction and anaerobic methodologies, with kinetic analyses and protein structure determination, to describe a previously uncharacterized enolate form of the chromophore that we propose is an intermediate in GFP fluorophore maturation. Isolation of this enolate intermediate will allow specific probing of the rate-limiting oxidation step for fluorophore biosynthesis.

GFP fluorophore biosynthesis is robust; variants with mutations at or near the chromophore, such as solubility-optimized GFPsol (F64L S65T F99S M153T V163A)[6] and blue fluorescent BFPsol (GFPsol Y66H H148G), undergo backbone cyclization, dehydration, and oxidation post-translational chemistry like that of the wild-type protein.[1,8] The oxidation reaction is rate-limiting, relatively slow ($t_{1/2}$ = 84 min^{-1}), and required to generate the visibly colored fluorophore of wild-type GFP.[2−5] A better understanding of this oxidation step may lead to the design of faster maturing mutant proteins with enhanced applications in cell labeling and biotechnology.[5]

Here we have chemically reduced[9] GFPsol and BFPsol with dithionite[4,10] to generate colorless samples that re-form fluorophores upon exposure to air. The GFPsol variant exhibited oxidation kinetics ($t_{1/2}$ = 39 min^{-1}, Figure 2) that are less than 2-fold different from those reported previously for the S65T variant ($t_{1/2}$ = 76 min^{-1}) by Reid and Flynn[4] (under somewhat different experimental conditions[11]). Reid and Flynn established that the chemically reduced moiety was kinetically competent to be an intermediate in fluorophore biosynthesis.[4] Interestingly, aerobic and anaerobic structures of the S65G Y66G GFP variant, which can cyclize the Gly−Gly−Gly sequence, indicated that formation and/or stabilization of the cyclized moiety is oxidation-dependent.[6] Thus, we anaerobically[9] crystallized and determined X-ray crystal structures[12]

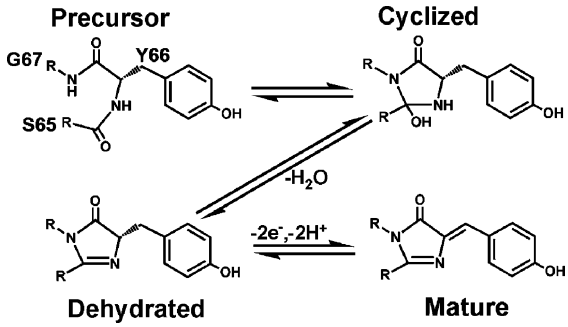

Figure 1. Proposed post-translational modifications for GFP chromophore biosynthesis.

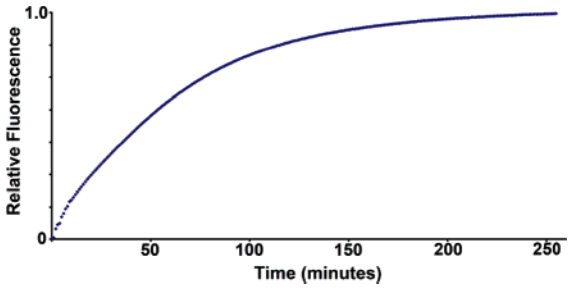

Figure 2. Kinetics of GFPsol chromophore oxidation monitored with fluorescence development. Dithionite-reduced GFPsol was brought out of the glovebox, exposed to air, and diluted into 100 mM Hepes pH 8.0 buffer. Fluorescence was measured at 25 °C via a Fluoromax-3 (Jobin-Yvon, Spex) fluorimeter using an excitation wavelength of 489 nm and emission wavelength at 508 nm. The fluorescence data were plotted (ln[A(max) − A(t)] vs time) to determine the pseudo-first-order rate constant, k_{obs}.

of reduced GFPsol and BFPsol to determine (1) the chemical nature of the reduced cyclized moiety, and (2) whether oxidation is required for ring stability.

In the 1.25 Å resolution structure of the dithionite-treated GFPsol (Figure 3A, Table 1), the imidazolone ring remained cyclized and dehydrated, but the chromophore was no longer constrained to be planar. The omit electron density revealed a mixture of chromophore conformations. We fit and refined the two best-defined conformers, which differ by a ∼10° tilt of their imidazolone rings and by side-chain rotations of Tyr66 (and neighboring His148) (Figure 3A). In the major conformer (62% occupancy), the Tyr66 phenolic oxygen formed hydrogen bonds with the Thr203 side-chain oxygen and His148 Nδ1 atoms. This conformer has nearly coplanar Tyr and imidazolone rings and resembles the mature chromophore. We argue against a significant population of mature chromophore, however, based on the ∼10° Tyr66 Cβ−Cγ rotation and, more importantly, the fact that the dithionite-treated protein and crystals are colorless.[13] In the minor chromophore conformer (38% occupancy), the imidazolone ring is tilted (away from Arg96) relative to the mature fluorophore, the His148 side chain fills part of the

10.1021/ja0552693 CCC: $33.50

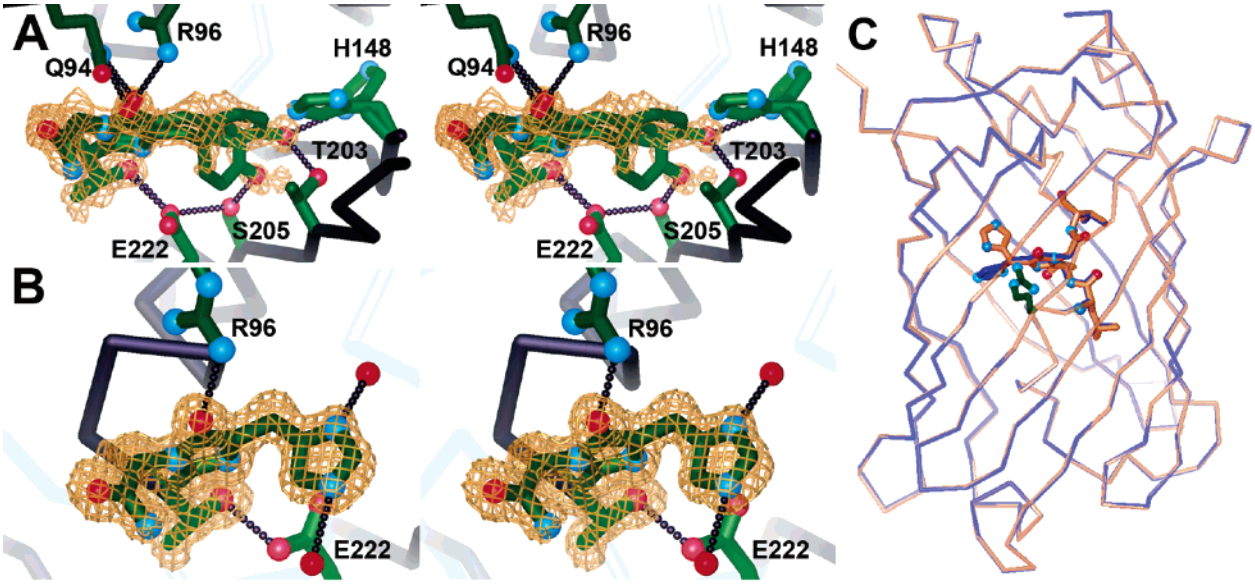

Figure 3. Crystallographic structures of dithionite-treated GFPsol and BFPsol variants. (A) Stereoimage for the 1.20 Å resolution GFPsol structure displayed with composite omit $2F_o-F_c$ electron density maps contoured at 1σ (gold). (B) Stereoimage for the 1.40 Å resolution BFPsol structure displayed with $2F_o-F_c$ electron density maps contoured at 1σ (gold). (C) Structural overlay of BFPsol before (yellow) and BFPms1 (PDB code 1KYP, blue)[14] after oxidation, highlighting the β-barrel fold of GFP. The chromophores and flanking residues, F64L and Val68, are displayed with Arg96 (green, below chromophore) for the two structures. Diffraction data sets were indexed and reduced with the HKL package,[19] and phases were determined by molecular replacement with AmoRe,[20] using previously refined GFP structures as search models.[1,14] Structures were manually fitted into difference electron density and omit maps with the XtalView package[21] and refined in SHELX-97[22] using all the diffraction data, except for 5% used for R_{free} calculations.[23] Standard uncertainties for atomic positions were determined by inverting the full least-squares covariance matrix in SHELX-97.[22]

Table 1. Crystallographic Data and Refinement Statistics

	GFPsol reduced	BFPsol reduced
resolution range (Å)	20−1.25	20−1.40
space group	$P2_12_12_1$	$P2_12_12_1$
unit cell (Å)	51.3, 62.5, 71.7	51.1, 62.5, 71.5
total/unique observations	154670/62613	157458/44523
completeness (%)	96.9 (91.5)[a]	97.3 (99.5)
R_{sym} (%)[b]	4.5 (34.9)	4.6 (35.5)
I/sigma	19.6 (2.3)	25.4 (3.2)
Wilson B-factor (Å^2)	9.9	14.4
refinement parameters	20543	19988
rmsd bond length (Å)	0.022	0.010
rmsd bond angle (Å)[c]	0.032	0.028
R-factor (R_{free}) (%)[d]	14.1 (20.0)	13.6 (19.8)
PDB code	2FZU	2FWQ

[a] Last shell in parentheses: GFPsol (1.29−1.25 Å) and BFPms1 (1.45−1.40 Å). [b] $R_{sym} = \sum |I_{hkl} - \langle I\rangle|/\sum \langle I\rangle$, where $\langle I\rangle$ is the average individual measurement of I_{hkl}. [c] Root mean square deviation bond angle is defined as the 1−3 bond distance (DANG restraint in SHELX-97).[22] [d] R-factor $= (\sum|F_o-F_c|)/\sum|F_o|$.

volume occupied by Tyr66 of the oxidized chromophore, and the rotated Tyr66 side-chain oxygen forms a hydrogen bond to the side-chain hydroxyl of Ser205 (Figure 3A). The imidazolone and phenolic ring planes are rotated ~60° from each other about the Tyr66 Cβ−Cγ bond, precluding conjugation of the two rings and indicating Cβ sp^3-hybridization and reduction of the Cα−Cβ double bond of the mature chromophore. The lack of a single conformer for reduced GFPsol is likely due to constraints imposed by the protein architecture, which complements the mature chromophore. Together, our GFPsol structural data are consistent with an sp^3-hybridized Cβ atom and reduction of the Cα−Cβ bond.

The 1.40 Å resolution structure (Table 1) of reduced BFPsol also revealed a cyclized, dehydrated imidazolone ring, but with a single, well-defined chromophore conformer. The nitrogen atoms of the Y66H imidazole side chain formed hydrogen bonds with a pair of ordered water molecules (Figure 3B). Compared to the planar mature chromophore,[8,14] the reduced Y66H side chain is rotated ~105° about the Cα−Cβ bond and flipped ~180° about the Cβ−Cγ bond, making the imidazole and imidazolone ring planes perpendicular (Figure 3C). In contrast, similar Cα−Cβ torsions for Tyr66 in the reduced GFPsol chromophore appear to be restricted by collisions with the protein scaffold. Conservation of the planar five-membered imidazolone ring in both the reduced GFPsol and reduced BFPsol structures demonstrates that oxidation is not required for ring stabilization. Importantly, the single conformer for reduced BFPsol and our high-resolution data and analyses allow the assignment of the reduced chromophore structure.

Our GFPsol and BFPsol high-resolution structural analyses support an enolate moiety for the dithionite-reduced chromophore. The planar BFPsol imidazolone ring and sp^2-hybridized Y66H Cα atoms indicate that the Cα atom must contain a double bond to the Cβ, backbone nitrogen, or carbonyl carbon atom. Our BFPsol structure precludes a Cα−Cβ double bond based on the Cα−Cβ bond length[15] (1.48 ± 0.05 Å) and the unambiguous sp^3-hybridization of the Cβ atom. We can also rule out a Cα−N double bond, which would require placement of a positively charged Gly67 nitrogen atom adjacent to positively charged Arg96. Instead, our BFPsol data support the assignment of a Cα−C double bond, concomitantly producing a negative charge on the enolate oxygen atom (Figure 4). An enolate assignment is also supported by the Y66H Cα to carbonyl carbon (1.36 ± 0.05 Å) and carbonyl carbon−oxygen (1.39 ± 0.05 Å) bond lengths (consistent with a C=C double and C−O single bond,[15] respectively) and the short (2.6 Å) hydrogen bond (salt bridge) between conserved Arg96 and the enolate oxygen atom (Figures 3B and 4). Our reduced GFPsol structure is also consistent with an enolate assignment. Thus, our structural data and analyses for GFP variants provide strong evidence for generation of an enolate moiety upon reduction of the mature chromophore.

Together, the analyses of our structural and kinetic data support this enolate moiety as an uncharacterized intermediate in GFP chromophore biosynthesis (Figure 4). We suggest that Arg96

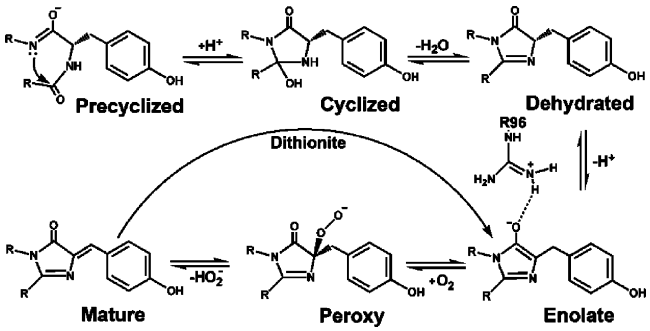

Figure 4. Proposed reaction scheme for GFP chromophore biosynthesis, including new enolate intermediate.

acidifies the Tyr66 Cα atom of the dehydrated intermediate to facilitate enolate intermediate formation. We and others have already established that Arg96 has an electrophilic role in backbone cyclization, acidifying the Gly67 amide nitrogen and/or Tyr66 Cα atoms to help drive ring formation.[6,16] Here we propose a similar deprotonation of the dehydrated intermediate in chromophore biosynthesis, to create a stable enolate intermediate with an aromatic π-system that is primed for Tyr66 Cα−Cβ oxidation and chromophore maturation.

In our proposed mechanism, molecular oxygen is incorporated on the Arg96 face of this enolate intermediate for rate-limiting oxidation to complete fluorophore biosynthesis (Figures 3 and 4). The tetrahedral Cα geometry of the resulting peroxy chromophore intermediate might be accommodated by the Tyr66 Cβ atom position and side-chain conformation observed in our minor reduced GFPsol conformer, combined with the imidazolone ring position of the mature chromophore (Figure 3A). In this scenario, conserved Arg96 is properly positioned to provide positive electrostatic interactions for chromophore oxidation. This mechanism has analogies to topaquinone (TPQ) cofactor biogenesis in amine oxidase.[17] In amine oxidase, molecular oxygen binds near a copper ion, which provides electrostatic interactions to accelerate electron transfer from the cofactor (generating a radical cation) to molecular oxygen. Recombination of the superoxide radical and radical cation leads to oxidation producing the mature cofactor. We hypothesize that GFP undergoes post-translational chemistry analogous to TPQ biosynthesis in amine oxidase. Positively charged Arg96 drives electron transfer from the enolate intermediate to oxygen, facilitating oxygen incorporation to generate a peroxy intermediate, followed by oxidation to generate a mature chromophore (Figure 4).

The chromophore bond lengths and geometries from these high-resolution crystallographic structures provide compelling evidence for a stable, reduced enolate form of the GFP chromophore. Together with previous results,[4] our kinetic analyses indicate that this enolate moiety is kinetically competent to be an intermediate in chromophore biosynthesis. Although our data cannot rule out other models for chromophore biosynthesis,[18] we propose that the simplest and most logical explanation for our data is an extension of the cyclization−dehydration−oxidation mechanism (Figure 4), originally proposed by Cubitt et al.[3] Importantly, we demonstrate that, by chemically reducing the mature chromophore, we effectively separate the cyclization and oxidation post-translational chemistry for GFP. This methodology allows testing the roles of specific residues, such as Arg96, in the rate-limiting oxidation reaction for fluorophore biosynthesis in GFP and its homologues. Such targeted characterizations may lead to the design of faster maturing proteins with applications in biotechnology and cell biology. Moreover, our results reveal how GFP mimics enzymes by creating a protein environment that stabilizes an otherwise high energy enolate intermediate to achieve its extraordinary protein self-modification.

Acknowledgment. We thank C. D. Putnam, T. I. Wood, and V. A. Roberts for discussions, and the La Jolla Interfaces in Sciences and NIH GM19290 Postdoctoral Fellowships (D.P.B.), NIH Grant GM37684, and the Stanford Synchrotron Radiation Laboratory.

References

(1) (a) Ormö, M.; Cubitt, A. B.; Kallio, K.; Gross, L. A.; Tsien, R. Y.; Remington, S. J. *Science* **1996**, *273*, 1392−1395. (b) Yang, F.; Moss, L. G.; Phillips, G. N., Jr. *Nat. Biotechnol.* **1996**, *14*, 1246−1251.
(2) Heim, R.; Prasher, D. C.; Tsien, R. Y. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 12501−12504.
(3) Cubitt, A. B.; Heim, R.; Adams, S. R.; Boyd, A. E.; Gross, L. A.; Tsien, R. Y. *Trends Biochem. Sci.* **1995**, *20*, 448−455.
(4) Reid, B. G.; Flynn, G. C. *Biochemistry* **1997**, *36*, 6786−6791.
(5) Heim, R.; Cubitt, A. B.; Tsien, R. Y. *Nature* **1995**, *373*, 663−664.
(6) Barondeau, D. P.; Putnam, C. D.; Kassmann, C. J.; Tainer, J. A.; Getzoff, E. D. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 12111−12116.
(7) Barondeau, D. P.; Kassmann, C. J.; Tainer, J. A.; Getzoff, E. D. *Biochemistry* **2005**, *44*, 1960−1970.
(8) Wachter, R. M.; King, B. A.; Heim, R.; Kallio, K.; Tsien, R. Y.; Boxer, S. G.; Remington, S. J. *Biochemistry* **1997**, *36*, 9759−9765.
(9) Mature GFPsol and BFPsol (ref 14) at 10 mg/mL were reduced with 5 mM dithionite (ref 10) for 1−2 days in an anaerobic glovebox. Chromophore reduction was monitored by loss of visible absorbance.
(10) Inouye, S.; Tsuji, F. I. *FEBS Lett.* **1994**, *351*, 211−214.
(11) The Reid and Flynn dithionite reduction process,[4] unlike ours, required protein denaturation by incubation at 95 °C in a 8 M urea and 1 mM DTT solution. Given this caveat and the slight differences in temperature (room temperature vs our 25 °C), pH (7.5 vs our 8.0), and mutation sets (we not only have the S65T mutation, but additional solubility mutations F64L, F99S, M153T, and V163A), our oxidation rates are similar.
(12) Reduced GFP proteins (10−15 mg/mL) were crystallized in hanging drops (refs 1 and 14). X-ray diffraction data were collected at the Stanford Synchrotron Radiation Laboratory on beamline 7-1 using a Mar345 image plate detector and a wavelength of 1.08 Å.
(13) The high GFPsol chromophore extinction coefficient (∼55 000 M^{-1} cm^{-1}) and protein concentration in our crystals (only 37% solvent) would make even minor populations of the mature chromophore observable.
(14) Barondeau, D. P.; Kassmann, C. J.; Tainer, J. A.; Getzoff, E. D. *J. Am. Chem. Soc.* **2002**, *124*, 3522−3524.
(15) Typical bond lengths for comparison: C−C 1.53 Å, C═C 1.34 Å, C−O 1.43 Å, and C═O 1.23 Å.
(16) (a) Sniegowski, J. A.; Lappe, J. W.; Patel, H. N.; Huffman, H. A.; Wachter, R. M. *J. Biol. Chem.* **2005**, *280*, 26248−26255. (b) Sniegowski, J. A.; Phail, M. E.; Wachter, R. M. *Biochem. Biophys. Res. Commun.* **2005**, *332*, 657−663. (c) Wood, T. I.; Barondeau, D. P.; Hitomi, C.; Kassmann, C. J.; Tainer, J. A.; Getzoff, E. D. *Biochemistry* **2005**, *44*, 16211−16220.
(17) (a) Goto, Y.; Klinman, J. P. *Biochemistry* **2002**, *41*, 13637−13643. (b) Su, Q.; Klinman, J. P. *Biochemistry* **1998**, *37*, 12513−12525.
(18) (a) Siegbahn, P. E. M.; Wirstam, M.; Zimmer, M. *Int. J. Quantum Chem.* **2001**, *81*, 169−186. (b) Rosenow, M. A.; Huffman, H. A.; Phail, M. E.; Wachter, R. M. *Biochemistry* **2004**, *43*, 4464−4472.
(19) Otwinowski, Z.; Minor, W. *Macromol. Crystallogr., Part A* **1997**, *276*, 307−326.
(20) Navaza, J. *Acta Crystallogr.* **1994**, *A50*, 157−163.
(21) McRee, D. E. *J. Struct. Biol.* **1999**, *125*, 156−165.
(22) Sheldrick, G.; Schneider, T. *Methods Enzymol.* **1997**, *277*, 319−343.
(23) Brunger, A. T. *Nature* **1992**, *355*, 472−474.

JA0552693