J|A|C|S
ARTICLES
Published on Web 03/17/2006

Reaction Progress of Chromophore Biogenesis in Green Fluorescent Protein

Liping Zhang, Hetal N. Patel, Jason W. Lappe, and Rebekka M. Wachter*

Contribution from the Department of Chemistry and Biochemistry, Arizona State University, Tempe, Arizona 85287-1604

Received November 26, 2005; E-mail: RWachter@asu.edu

Abstract: The mature form of green fluorescent protein (GFP) is generated by a spontaneous self-modification process that is essentially irreversible. A key step in chromophore biosynthesis involves slow air oxidation of an intermediate species, in which the backbone atoms of residues 65−67 have condensed to form a five-membered heterocycle. We have investigated the kinetics of hydrogen peroxide evolution during in vitro GFP maturation and found that the H_2O_2 coproduct is generated prior to the acquisition of green fluorescence at a stoichiometry of 1:1 (peroxide/chromophore). The experimental progress curves were computer-fitted to a three-step mechanism, in which the first step proceeds with a time constant of 1.5 (±1.1) min and includes protein folding and peptide cyclization. Kinetic data obtained by HPLC analysis support a rapid cyclization reaction that can be reversed upon acid denaturation. The second step proceeds with a time constant of 34.0 (±1.5) min and entails rate-limiting protein oxidation, as supported by a mass loss of 2 Da observed for tryptic peptides derived from species that accumulate during the reaction. The final step in GFP maturation proceeds with a time constant of 10.6 (±1.2) min, suggesting that this step may contribute to overall rate retardation. We propose that under highly aerobic conditions, the dominant reaction path follows a cyclization−oxidation−dehydration mechanism, in which dehydration of the heterocycle is facilitated by slow proton abstraction from the Tyr66 β-carbon. In combination, the results presented here suggest a role for molecular oxygen in trapping the cyclized form of GFP.

Introduction

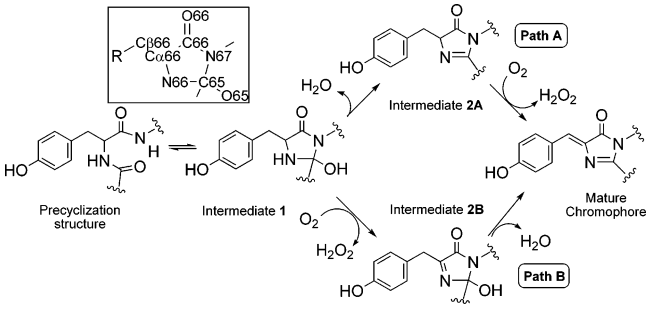
Oxidative tyrosine modifications using molecular oxygen as electron acceptor have been characterized in a number of post-translational protein modifications. Some of these reactions are spontaneous protein processing events that require redox-active metal ions. For example, in the copper amine oxidases, the phenolic ring of a tyrosine residue is oxidized to generate the built-in cofactor topaquinone,[1] and in galactose oxidase, a thioether bridge connects a tyrosine phenolic ring to a nearby cysteine residue.[2] Other kinds of oxidative tyrosine modifications occur in the absence of metal ions and require enzyme systems, such as the biosynthesis of pyrroloquinoline quinone in the bacterial dehydrogenases.[3] In green fluorescent protein (GFP), chromophore formation is autocatalytic and does not require metal ions. The O_2-dependent processing event involves the oxidation of a tyrosine residue along the Cα−Cβ bond and is triggered by a spontaneous peptide backbone cyclization reaction that occurs in the protein's interior once the correct tertiary structure is attained. Upon completion of the maturation process, GFP bears an amino acid-derived covalently bound chromophore that is bright green fluorescent.[4] For this reason, GFP represents a genetically encodable fluorophore and is commonly used as a reporter molecule in biotechnological and cellular imaging applications. However, several mechanistic aspects of the protein self-modification process continue to be debated.

GFP chromophore formation spontaneously follows protein folding. The protein's 11-stranded β-barrel surrounds an internal helix that bears the chromophore-forming residues Ser65, Tyr66, and Gly67 (wild-type GFP). Nucleophilic attack of the amide nitrogen of Gly67 onto the carbonyl carbon of Ser65 leads to intrachain ring closure (Scheme 1), generating an intermediate with a five-membered ring formed from the backbone atoms of the original polypeptide chain.[5,6] Full chromophore maturation, and hence the development of visible fluorescence, requires further processing steps that consist of oxidation of a single to a double bond, as well as the elimination of water from the heterocyclic intermediate (Scheme 1). According to the original mechanistic proposal,[6] dehydration immediately follows ring closure (Path A, Scheme 1), thus stabilizing the resulting imidazolinone ring.[4] Under aerobic conditions, the C_α66−C_β-66 single bond (for atom labels, see Scheme 1 inset) is oxidized to attain π-conjugation of the phenolic group with the imidazolinone ring. This mechanism was proposed on the basis of live cell experiments that supported the accumulation of a stable GFP intermediate in an inert atmosphere.[5,6] *Escherichia coli*

(1) DuBois, J. L.; Klinman, J. P. *Biochemistry* **2005**, *44*, 11381−11388.
(2) Firbank, S.; Rogers, M.; Guerrero, R. H.; Dooley, D. M.; Halcrow, M. A.; Phillips, S. E.; Knowles, P. F.; McPherson, M. J. *Biochem. Soc. Symp.* **2004**, *71*, 15−25.
(3) Magnusson, O. T.; Toyama, H.; Saeki, M.; Rojas, A.; Reed, J. C.; Liddington, R. C.; Klinman, J. P.; Schwarzenbacher, R. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 7913−7918.
(4) Tsien, R. Y. *Annu. Rev. Biochem.* **1998**, *67*, 509−544.
(5) Heim, R.; Prasher, D. C.; Tsien, R. Y. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 12501−12504.
(6) Cubitt, A. B.; Heim, R.; Adams, S. R.; Boyd, A. E.; Gross, L. A.; Tsien, R. Y. *Trends Biochem. Sci.* **1995**, *20*, 448−455.

10.1021/ja0580439 CCC: $33.50 **© 2006 American Chemical Society**

Scheme 1. Proposed Reaction Scheme for GFP Chromophore Biogenesis (Inset: Atom Labels Used in the Text)



cultures expressing GFP were first grown in the absence of molecular oxygen to ensure complete conversion of the GFP precursor to an intermediate form.[5] After 3 days, cells were exposed to air to allow for maturation to the green stage, while fresh protein synthesis was inhibited. The switch from an inert to an aerobic atmosphere was reported to result in a reduction of mass by 1 (±4) Da,[6] consistent with protein oxidation via dehydrogenation (−2 Da) as the last step in chromophore formation. These data, derived from in vivo maturation experiments, form the foundation for Path A (Scheme 1). The protein oxidation step suggests the formation of hydrogen peroxide as the O_2 reduction product,[4] though the reduced oxygen species has not been identified yet.

Recently, we have considered the possibility that GFP may mature via an alternate pathway, Path B (Scheme 1), in which the dehydration and oxidation steps are essentially reversed. X-ray crystallographic and spectroscopic data on the Y66L variant have suggested that the condensation product is trapped via ring oxidation, in accord with a reaction pathway where subsequent dehydration leads to the Y66L analogue of the GFP chromophore.[7] According to Path B, the backbone condenses to form a cyclic tetrahedral intermediate, and the ring is subsequently oxidized to the cyclic imine by molecular oxygen. In the third and final step, elimination of the hydroxyl leaving group as a water molecule is coupled to a proton-transfer reaction that may proceed via hydrogen-bonded solvent molecules.[8] Though replacement of the aromatic Tyr66 with an aliphatic residue appears to reduce the efficiency of ring dehydration, an aromatic residue is not required for cyclization or oxidation, as further substantiated by GFP variants that contain the Y66G or the Y66S substitution.[9,10]

The rate of GFP maturation has been measured both in vivo and in vitro. Time constants for fluorescence acquisition were determined after admission of air to GFP-expressing bacterial cells or freshly prepared cell lysates and were reported to be 120 min for wild-type GFP and 27 min for GFP-S65T.[5,11] The data fit well to a single exponential and suggested that the oxidation step may be rate-determining in the overall chromophore formation process.[4] Surprisingly, in vitro maturation of GFP-S65T using inclusion bodies as starting material was reported to proceed with a time constant of 122 min,[12] four times more slowly than in live cells.[11] More recently, we have determined the de novo maturation rate of EGFP (GFP-F64L/S65T) to be 60 min under very similar conditions.[13] The rate of fluorescence reacquisition has also been determined after chemical reduction of the mature chromophore, which is reversible upon exposure to molecular oxygen.[14] In vitro reoxidation rates using purified GFP-S65T were reported to give time constants on the order of 108 min,[12] similar to the reported de novo maturation (122 min). Surprisingly fast reoxidation rates such as $\tau = 2$ min have recently been reported for some novel yellow variants.[15] However, it is unclear in what way reoxidation rates are related to de novo chromophore biosynthesis.

The rate of the main-chain condensation reaction has not been measured yet directly. Some years ago, the cyclization time constant was estimated to be less than 4 or 5 min for GFP-S65T,[12] as inferred from other kinetic data assuming a sequential mechanism. Recently, backbone ring closure has been suggested to be a thermodynamically unfavorable process that may lead to a high-energy unstable intermediate.[9] This intermediate may be trapped either via subsequent dehydration[10] or via ring oxidation.[8] Several X-ray structures of precyclization states in GFP mutants confirmed that the chromophore-forming tripeptide is locked into a tight-turn conformation that facilitates backbone cross-linking.[9] The highly conserved residues Arg96 and Glu222 positioned close to the chromophore in the tertiary structure have been shown to play a major role in the backbone cyclization kinetics.[9,13,16] These data support a model for GFP chromophore synthesis in which the carboxylate of Glu222 facilitates proton abstraction from the Gly67 amide nitrogen or the Tyr66 α-carbon, whereas Arg96 plays the role of an electrophile by lowering the respective pK_a values and stabilizing the α-enolate.

Here, we have examined the kinetics of in vitro GFP maturation in relation to the evolution of hydrogen peroxide. We have characterized the molecular mass of an oxidized intermediate and found that, in air-saturated aqueous solutions, a dehydrogenated species is generated. These observations have important implications in determining the parameters that limit the GFP maturation rate.

Results and Discussion

During GFP Maturation, Hydrogen Peroxide Is Produced at a 1:1 Molar Ratio. To investigate whether protein oxidation plays a role in trapping the heterocyclic intermediate formed during GFP maturation, we monitored the in vitro kinetics of chromophore formation in relation to dioxygen reduction. Immature GFP was derived from inclusion bodies, which consist of insoluble misfolded protein that accumulates in the *E. coli* cell upon overexpression at 42 °C.[12] His-tagged GFP inclusion bodies were washed extensively to remove native GFP and other protein contaminants, solubilized in 8 M urea and purified over a Ni-NTA affinity column under denaturing conditions.[8] De novo protein folding of freshly purified material (same day)

(7) Rosenow, M. A.; Patel, H. N.; Wachter, R. M. *Biochemistry* **2005**, *44*, 8303−8311.
(8) Rosenow, M. A.; Huffman, H. A.; Phail, M. E.; Wachter, R. M. *Biochemistry* **2004**, *43*, 4464−4472.
(9) 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.
(10) Barondeau, D. P.; Kassmann, C. J.; Tainer, J. A.; Getzoff, E. D. *Biochemistry* **2005**, *44*, 1960−1970.
(11) Heim, R.; Cubitt, A. B.; Tsien, R. Y. *Nature* **1995**, *373*, 663−664.
(12) Reid, B. G.; Flynn, G. C. *Biochemistry* **1997**, *36*, 6786−6791.
(13) Sniegowski, J. A.; Lappe, J. W.; Patel, H. N.; Huffman, H. A.; Wachter, R. M. *J. Biol. Chem.* **2005**, *280*, 26248−26255.
(14) Inouye, S.; Tsuji, F. I. *FEBS Lett.* **1994**, *351*, 211−214.
(15) Nagai, T.; Ibata, K.; Park, E. S.; Kubota, M.; Mikoshiba, K.; Miyawaki, A. *Nat. Biotechnol.* **2002**, *20*, 87−90.
(16) Sniegowski, J. A.; Phail, M. E.; Wachter, R. M. *Biochem. Biophys. Res. Commun.* **2005**, *332*, 657−663.

was induced by rapid dilution into folding buffer, and the time course of maturation was monitored for 2.5 h. All experiments were performed using inclusion bodies of GFP-trix, a variant prepared in our laboratory that incorporates the mutation S65T known to enhance the intensity of green light emission,[11] the substitutions F64L and F99S/M153T/V163A known to improve in vivo folding,[17,18] and the substitution A206K known to eliminate protein dimerization.[19]

The reduced oxygen species generated during chromophore formation was determined to be hydrogen peroxide, which evolved at a 1:1 molar ratio of H_2O_2/chromophore. Hydrogen peroxide concentrations were monitored by an enzyme-linked colorimetric assay based on Amplex Red (10-acetyl-10*H*-phenoxazine-3,7-diol) and horseradish peroxidase (HRP), in which the reagent was oxidized to resorufin as a function of H_2O_2 at a 1:1 molar ratio of resorufin to H_2O_2.[20] At the end point of GFP maturation (~2.5 h), the resorufin concentration was determined by absorbance at 571 nm, and the chromophore concentration was determined by absorbance at 489 nm (ϵ_{489} = 33 787 M^{-1} cm^{-1} for GFP-trix). In six independent experiments, the average molar ratio of resorufin/chromophore was determined to be 1.043 (±0.139, $n = 6$), with sample standard deviation in parentheses. To examine whether protein denaturation would result in the release of additional peroxide,[3] aliquots of the maturation reaction were acid-denatured, incubated for 10 min, and then neutralized immediately before addition of the Amplex Red reagent. Using this method, we determined the average molar ratio of resorufin/chromophore to be 1.050 (±0.154, $n = 3$), essentially unmodified by protein denaturation. We conclude that one molecule of hydrogen peroxide is generated per mature GFP and that hydrogen peroxide is readily released into solution from the protein's interior.

The Amplex Red assay cannot be used to discriminate between the two potential coproducts of the reaction, superoxide and hydrogen peroxide,[4] since superoxide is known to rapidly disproportionate to O_2 and H_2O_2. However, direct hydrogen peroxide formation, possibly via a hydroperoxy adduct to the heterocycle,[8] appears more likely due to the strictly maintained molar ratio of one peroxide molecule per GFP chromophore.

Hydrogen Peroxide Evolution Precedes the Acquisition of Green Fluorescence. To determine the kinetics of hydrogen peroxide evolution during the process of chromophore biogenesis, a continuous fluorescence assay was employed. Protein folding was induced by rapid dilution of inclusion body-derived protein as described above. The reaction was carried out in a stirred fluorimeter cell containing both folding buffer and Amplex Red reagent. The enzyme-catalyzed production of resorufin as a function of H_2O_2 was monitored continuously by resorufin fluorescence at 580 nm (Figure 1A). Since the production of the reporter molecule from H_2O_2 introduces a significant time delay, the rate of this reaction was determined independently by the use of three standard H_2O_2 solutions (0.1−0.3 μM H_2O_2, no GFP) (Figure 2). Assuming pseudo-first-order kinetics due to excess reagent, the rate constant for the conversion of H_2O_2 to resorufin was extracted to be 0.114 (±0.016, $n = 3$) min^{-1} by computer-fitting the data to an equation of the form $F = F_{max} - e^{(-kt)} \cdot F_{max}$, where F equals fluorescence emission intensity.

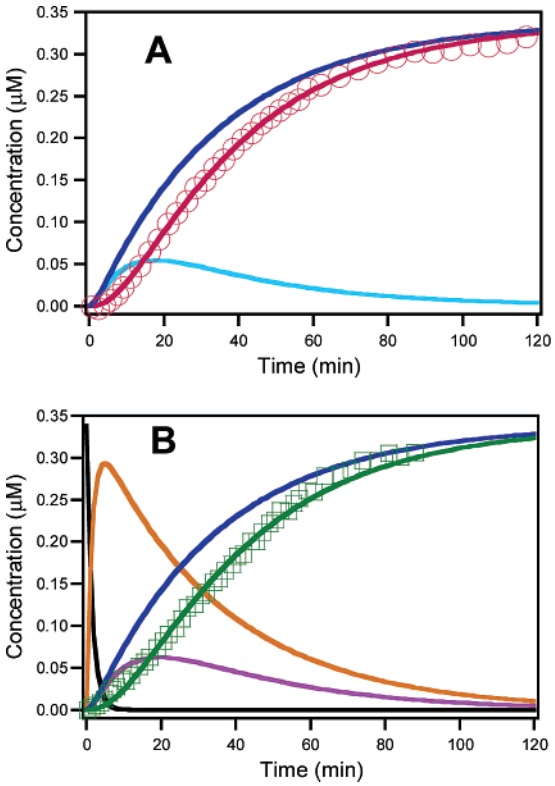


Figure 1. Progress curves for GFP maturation. Experimentally determined resorufin and GFP chromophore concentrations as a function of time (symbols), and progress curves calculated from the rate constants extracted by DynaFit curve-fitting procedures (solid lines). (A) Experimental resorufin concentration, red circles (20% of collected data shown). Calculated progress curves: Resorufin, red; free hydrogen peroxide, light blue; total hydrogen peroxide, dark blue. (B) Experimental chromophore concentration, green squares (30% of collected data shown). Calculated progress curves: GFP precursor, black; intermediate 1, orange; intermediate 2, purple; total hydrogen peroxide, dark blue; mature chromophore, green.

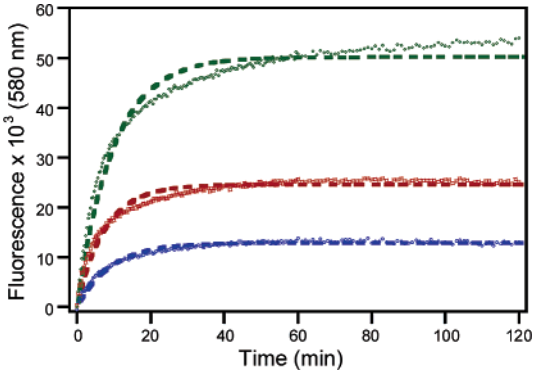
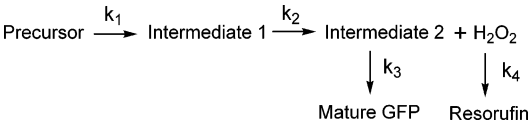


Figure 2. Response curves of standard hydrogen peroxide solutions in the presence of Amplex Red reagent (baseline-corrected). The fluorescence of the oxidation product resorufin is monitored via emission at 580 nm. Blue, 0.10 μM H_2O_2; red, 0.20 μM H_2O_2; green, 0.30 μM H_2O_2. The reaction is assumed to proceed via pseudo-first-order kinetics, and the data are fitted to a first-order rate equation as indicated by the dashed lines.

The acquisition of green fluorescence during GFP maturation was determined in a separate experiment by monitoring the chromophore emission at 510 nm (λ_{ex} = 480 nm) under identical conditions, but in the absence of Amplex Red reagent due to potential fluorescence resonance energy transfer between the

(17) Cormack, B. P.; Valdivia, R. H.; Falkow, S. *Gene* **1996**, *173*, 33−38.
(18) Crameri, A.; Whitehorn, E. A.; Tate, E.; Stemmer, W. P. C. *Nat. Biotechnol.* **1996**, *14*, 315−319.
(19) Zacharias, D. A.; Violin, J. D.; Newton, A. C.; Tsien, R. Y. *Science* **2002**, *296*, 913−916.
(20) Zhou, M.; Diwu, Z.; Panchuk-Voloshina, N.; Haugland, R. P. *Anal. Biochem.* **1997**, *253*, 162−168.

Scheme 2. Mechanistic Scheme Used for Global Curve Fitting of Progress Curves Obtained by Monitoring the Fluorescence of Resorufin and Mature GFP (k_1, k_2, k_3 Variables, k_4 Fixed)



GFP chromophore and resorufin. To determine whether the Amplex Red reagent itself has any effect upon the GFP maturation reaction, the reaction was quantized by reverse-phase HPLC (see below) with and without Amplex Red, and no interference was detected. Upon completion of maturation, the GFP chromophore concentration was determined by absorbance at 489 nm and was used to convert fluorescence emission values to chromophore concentration values for the entire maturation time course (Figure 1B).

For the analysis of kinetic data, the program DynaFit was used to computer-fit the experimentally determined progress curves for resorufin and GFP chromophore production to different kinetic schemes.[21] GFP chromophore biogenesis may be described as a sequential three-step process (rate constants k_1, k_2, and k_3) involving four protein species and the production of peroxide, which is converted to resorufin in a fourth step (rate constant k_4). Therefore, the time evolution of each of six species was modeled by using a set of six simultaneous differential equations (Scheme 2). Global curve fitting of fluorescence emission data was carried out for three different mechanisms, in which hydrogen peroxide release was modeled concomitantly with the first, the second, or the third step of chromophore formation. In the curve-fitting procedures, k_1, k_2, and k_3 were fitted parameters, whereas k_4 was fixed to its previously determined value of 0.114 ($\pm$0.016) min^{-1}. Model discrimination analysis using the Akaike Information Criterion (AIC),[22] as implemented in the program DynaFit, provides substantial support for a mechanistic model in which hydrogen peroxide is released in the second step as shown in Scheme 2 (AIC delta value = 0). For this model, an excellent fit to the experimental data is obtained. We have chosen to model the reaction with k_1 fast and k_2 slow, not vice versa (which results in identical fitting statistics), because oxidation has been shown to be rate-determining in GFP maturation.[5,6]

On the other hand, H_2O_2 production modeled as either part of the first[23] or the last step of chromophore formation results in unreasonably large standard errors on fitted parameters. These models are not supported by Akaike analysis (AIC δ value = 85 or 892, respectively). In addition, recent GFP crystal structures have established that cyclization, not oxidation, is the first chemical event in chromophore formation.[6,9]

According to the interpretation presented in Scheme 2, the first step in the reaction includes both protein folding and peptide backbone cyclization. The computer-fitted rate constant k_1 was determined to be 0.69 min^{-1} for the combined process, with an estimated error of $\pm$0.52 min^{-1} (this error estimate includes the error on k_4; see Experimental Section). The rate constant k_2 for generation of intermediate 2 and hydrogen peroxide was extracted to be 0.029 min^{-1}, estimated error $\pm$0.001 min^{-1}, and the rate constant k_3 for generation of the mature chromophore from intermediate 2 was extracted to be 0.094 min^{-1}, estimated error $\pm$0.011. In a sequential process, the overall transit time ($\tau = k^{-1}$) can be approximated by addition of the transit times for each individual process. Therefore, $\tau_{Total} = \tau_1 + \tau_2 + \tau_3 =$ 1.5 ($\pm$1.2) min + 34.0 ($\pm$1.5) min + 10.6 min ($\pm$1.2) = 46.0 ($\pm$2.2) min, somewhat more rapid than the time constant for EGFP maturation, which was determined to be 57 min by a monoexponential curve fit.[13] The increase in rate could be due to the improved folding efficiency of GFP-trix.[24]

The Ejection of Water from the Heterocycle May Contribute to Rate Retardation. During in vitro GFP maturation, a time delay is observed between the production of H_2O_2 and the acquisition of green fluorescence (Figure 1B). Free H_2O_2 generated during the protein self-modification process is continuously depleted by its reaction with Amplex Red to produce the reporter molecule resorufin ($\tau_4 = 8.77 \pm 1.23$ min). Therefore, the numerical addition of resorufin concentration to free hydrogen peroxide concentration, as calculated from the extracted rate constants, will yield a combined progress curve for total H_2O_2 produced by GFP during the reaction (Figure 1A,B). Comparison of the combined H_2O_2 progress curve with the development of green fluorescence indicates that the final step in protein maturation introduces an additional time delay of $\tau_3 = 10.6$ min. We propose that this last step consists of proton shuttling events that may include slow removal of a proton from the β-carbon of residue 66 (Cβ66) (Scheme 1 inset). A net proton transfer from Cβ66 to the hydroxyl adduct of the heterocycle (O65) may be essential in the elimination of water to yield the extended π-orbital system of the mature chromophore (Scheme 1).[8] If so, the rate of dehydration would be a function of the acidity of Cβ66, as proposed previously based on inefficient dehydration of Y66L.[7]

Internal Peptide Cyclization Appears To Be Rapid and Reversible upon Acid Denaturation. To better monitor the reaction progress of correctly folded GFP toward synthesizing the chromophore, we have developed an HPLC method to separate precyclization from postcyclization protein species. Upon rapid dilution from urea, 85−95% of total inclusion body-derived protein folds into a soluble form. However, HPLC assays indicate that only about 60% of the soluble population attains the green fluorescent state (compared to 25% maturation determined for EGFP under identical conditions). Tryptophan fluorescence experiments suggest that the remaining protein is unable to mature due to irreversible protein misfolding events (see Supporting Information). On reverse-phase HPLC, different forms of GFP are separated into a late-eluting peak associated with precursor states and an early-eluting peak associated with intermediate and mature states (Figure 3). The chromophore-bearing protein coelutes with colorless intermediates that convert to the mature form over time. Control injections of mature GFP before and after urea denaturation, as well as inclusion body-derived denatured GFP, suggest that the late-eluting peak contains only precyclization forms of the protein, whereas the early-eluting peak contains only postcyclization forms.

Within the 3.3 min dead time of the experiment, a large fraction of the precursor peak appears to be converted to a postcyclization form of the protein (Figure 4). This observation

(21) Kuzmic, P. *Anal. Biochem.* **1996**, *237*, 260−273.
(22) Burnham, K. P.; Anderson, D. R. *Model Selection and Multimodel Interference*; Springer-Verlag: New York, 2002.
(23) Siegbahn, P. E. M.; Wirstam, M.; Zimmer, M. *Int. J. Quantum Chem.* **2001**, *81*, 169−186.
(24) Remington, S. J. *Nat. Biotechnol.* **2002**, *20*, 28−29.

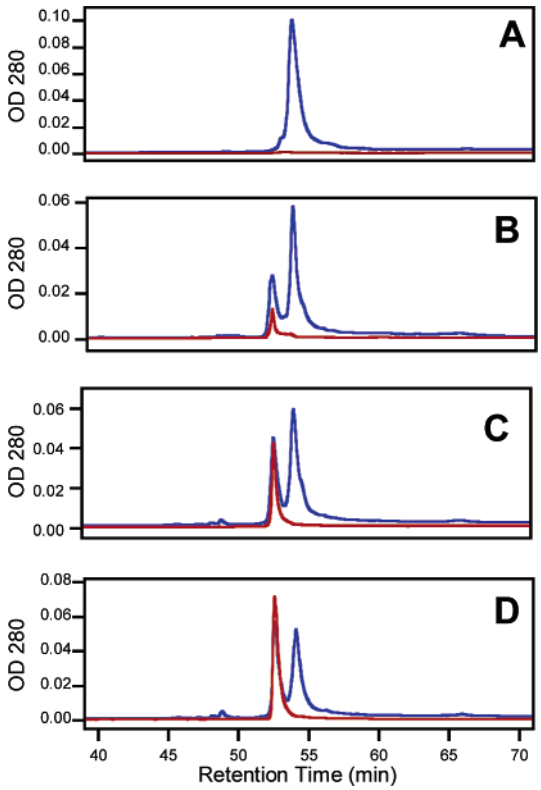


Figure 3. HPLC traces for samples collected at various time points of GFP maturation. (A) 0 min. (B) 6 min. (C) 35 min. (D) 100 min. Blue, absorbance monitored at 280 nm; red, absorbance monitored at 380 nm (GFP chromophore).

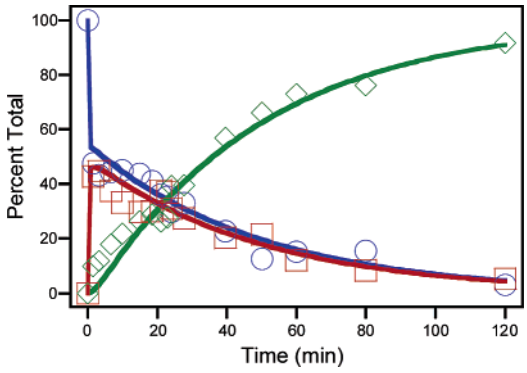


Figure 4. Precursor (blue), intermediates (red), and mature GFP (green) as a function of maturation time, as quantitated by reverse-phase HPLC. The data have been normalized to consider only the protein population that reaches the fully mature green fluorescent state. Due to on-column acid denaturation, the percent contribution of precyclization species (blue) and postcyclization intermediates (red) may be modified from the populations present in the intact protein during maturation. Therefore, the solid lines are drawn through the data to provide visual guidance only. However, GFP chromophore content (green) is not modified by acid denaturation, and hence accurately reflects the internal maturation process.

is consistent with rapid folding and backbone condensation, in accord with an estimated time constant of 1.5 ± 1.1 min for this process. On the basis of the chromophore absorbance at 380 nm under acidic conditions (Figure 3), the postcyclization peak can be quantitatively divided into contributions from colorless intermediate species and mature protein. Acid denaturation upon injection into the HPLC system may partially reverse ring closure, providing a rationale for the roughly equal proportions of precursor and intermediates observed throughout the reaction (Figure 4).

A Mass Loss of 2 Da Is Observed for an Intermediate that Accumulates during GFP Maturation. The progress curves constructed on the basis of curve fitting procedures imply that an oxidized intermediate may accumulate during the reaction (Figure 1B). To probe for intermediates with altered masses, early- and late-eluting HPLC peaks were collected for several maturation time points, lyophilized, and digested with trypsin. The resulting peptides were purified by reverse-phase HPLC. Masses of all eluting peptides were determined by matrix-assisted laser desorption ionization mass spectrometry (MALDI-MS) and matched with calculated masses obtained from theoretical tryptic digests of GFP-trix. Two peptides were observed that bear the chromophore-forming residues 65−67: peptide A spanning residues 53 to 73 with a theoretical mass of 2378.26 Da, and peptide B spanning residues 46 to 73 with a theoretical mass of 3128.63 (Table 1). For some of these peptides, a mass loss of 2 or 20 Da was observed (retention time 49−53 min), whereas for others, no mass loss was observed (retention time 55−57 min) (Table 1).

For a total of 14 mass determinations on peptides A and B derived from the precyclization HPLC peak (Figure 3), the average mass difference from the calculated mass was 0.06 (±0.39) Da, in accord with the presence of unmodified precursor protein only (Table 1). However, peptides A and B derived from the postcyclization HPLC peak suggest that a variety of protein species coelute. Peptides were observed with no mass change, as indicated by an average Δ mass of 0.13 (±0.26, $n = 8$) Da, and peptides were observed with a mass loss of 2 Da, as indicated by an average mass loss of −2.08 (±0.34, $n = 12$) Da. In addition, peptides were observed with a mass loss of 20 Da, as indicated by an average mass loss of −20.06 (±0.29, $n = 10$) Da. These data are in accord with a cyclization reaction that leads to the tetrahedral intermediate 1 (no mass loss), followed by the net loss of two hydrogen atoms due to slow oxidation (−2 Da), followed by the ejection of water to yield the mature chromophore (−20 Da).

The MALDI data provide additional support to a kinetic model in which ring closure is followed by air oxidation (−2 Da) to yield intermediate 2B (Scheme 1). Oxidative modifications by the sample preparation procedure are unlikely due to the large number of control experiments carried out in our laboratory.[13] In previous work, analogous peptides derived from mature GFP, immature R96M, mature and partially mature E222Q, and colorless Y66L were analyzed using the same methodology. In all instances, mature protein yielded peptides with a 20 Da mass loss, immature protein yielded peptides with no mass loss, and the Y66L variant, which contains a cyclized and oxidized backbone,[8] yielded peptides with a 2 Da mass loss. A mass loss of 18 Da, consistent with loss of water in the absence of oxidation (Path A), has not yet been observed in our laboratory. However, the hydration/dehydration equilibrium[7] in the protein's interior may be altered by the sample preparation procedure. For this reason, MALDI experiments cannot be used to exclude the possibility that a small population of the protein partitions through Path A.

Does the Reaction Pathway Depend on Oxygen Concentration? The kinetic data presented in this work provide support for Path B (Scheme 1) as the major in vitro GFP maturation mechanism in a stirred cell equilibrated with air.[1] The concentration of dissolved O_2 in the reaction chamber is estimated to

Table 1. MALDI Masses of Tryptic Peptides[a,b] Derived from Pre- and Postcyclization Species of GFP Isolated by HPLC at Maturation Time Points 3.3, 6.0, 35, and 100 min

	precyclization				postcyclization			
peptide[a,b]	mean	standard deviation	n^c	Δm^d	mean	standard deviation	n^c	Δm^d
A	2378.32	0.34	10	+0.06	2378.18	0.25	5	−0.08
A	−	−	−	−	2376.22	0.35	9	−2.04
A	−	−	−	−	2358.14	0.26	7	−20.12
B	3128.69	0.53	4	−0.06	3128.43	0.32	3	−0.20
B	−	−	−	−	3126.44	0.32	3	−2.19
B	−	−	−	−	3108.70	0.36	3	−19.94

[a] Amino acid sequence of tryptic peptide A: LPVPWPTLVTTLTYGVQCFSR (residues 53−73). Calculated mass: 2378.26 Da. [b] Amino acid sequence of tryptic peptide B: FICTTGKLPVPWPTLVTTLTYGVQCFSR (residues 46−73). Calculated mass: 3128.63. [c] Number of independent determinations. [d] Change in mass (experimental mean mass − calculated mass).

be 213 μM after correction for temperature and salinity.[25,26] Since data were obtained at one O_2 concentration only, a branched pathway (Scheme 1) cannot be excluded. If competition between oxidation and other chemical steps is modulated by the availability of molecular oxygen, Path A could become more prevalent upon oxygen depletion. In the intracellular milieu, oxygen levels may vary as a function of the cellular metabolic rate and oxygen diffusion through the growth medium. Dioxygen concentrations in live rat glioma cells were determined to be about 250 μM when cells were suspended in air-saturated saline,[27] similar to data obtained near single cells of mouse pancreatic islets.[28] However, O_2 concentrations were shown to decrease significantly as a function of glucose metabolism. Pericellular oxygen depletion as a result of respiration has also been demonstrated during ordinary tissue culture of mammalian cell lines.[29] Further studies are necessary to gain a better understanding of the oxygen dependence of GFP maturation and to be able to predict the rate of fluorescence acquisition in different cellular environments.

Conclusions

Protein oxidation is part of the mechanism of GFP maturation and is ultimately responsible for desaturation of the β-methylene bridge that ties the phenolic part of the chromophore to its heterocyclic end (Scheme 1). This process involves the net transfer of two hydrogen atoms from the protein to molecular oxygen and generates one hydrogen peroxide molecule per mature GFP chromophore. The in vitro data presented here are best interpreted in terms of a cyclization−oxidation−dehydration mechanism that proceeds with an overall time constant of 46 min in GFP-trix. Slow hydrogen peroxide evolution follows fast peptide backbone cyclization, which in turn is triggered by protein folding. Proton shuttling events are proposed to complete the overall process, culminating in the ejection of water from the heterocycle to yield the extended π-system of the mature chromophore (Scheme 1). This mechanism differs from that proposed some years ago based on live cell experiments,[6] in that the cyclic intermediate is proposed to be stabilized by oxidation rather than dehydration.

In the work presented here, kinetic data were collected on a variant termed GFP-trix, which contains the S65T substitution known to accelerate the GFP maturation rate.[4] The in vitro maturation process of GFP-trix is reminiscent of EGFP−Y66L,[8] where self-modification has been demonstrated to proceed via a cyclization−oxidation−dehydration mechanism, although with altered reaction energetics.[7] X-ray crystallographic data on EGFP−Y66L indicate that dehydration of the heterocycle is favored only upon extensive π-conjugation with the side chain-derived adduct. In this variant, the oxidized form of the protein is initially completely hydrated.[8] Subsequent slow enamine formation, which places a double-bond between C_α66 and C_β-66, results in partial dehydration of the ring. In the structure of the aerobic form of S65G/Y66G, the internal helix was reported to be cyclized to a hydroxylated imidazolidine dione, an entity derived from oxidation of the tetrahedral intermediate generated by ring closure.[9] Remarkably, under anaerobic conditions, the same variant was found to remain in the precyclization state, despite the close proximity of ring-forming atoms,[9] suggesting that oxygen may be important in stabilizing the cyclic form.

Experimental Section

Preparation and Purification of GFP-trix Derived from Inclusion Bodies. The GFP-trix variant (GFP− F64L/S65T/F99S/M153T/V163A/A206K) was prepared by site-directed mutagenesis using the QuikChange kit (Stratagene), using EGFP (GFP−F64L/S65T) as a parent clone. GFP-trix, N-terminally tagged with a six-His tag consisting of a total of 13 amino acid residues, was overexpressed in the $pRSET_B$ plasmid at 42 °C in *E. coli* strain JM109(DE3).[8] Under these conditions, most of the overexpressed GFP protein does not reach its native tertiary fold and does not form a chromophore, but instead partitions into inclusion bodies.[12] The inclusion bodies were washed exactly as described previously.[8] A total of five washes were performed, followed by solubilization in 8 M urea, 50 mM HEPES pH 7.9, 50 mM NaCl, and 1 mM DTT, and purification over a Ni-NTA affinity column as described.[8] By SDS-PAGE, the purified urea-denatured protein pool was estimated to be 90−95% homogeneous. The absence of a GFP chromophore was verified by HPLC (see below) by monitoring the absorbance at 380 nm, which is the absorbance maximum of the chromophore under denaturing acidic conditions.[30] HPLC was also used to estimate the protein concentration to be 0.23 mg/mL.

Protein Maturation Monitored via Hydrogen Peroxide Evolution. Hydrogen peroxide assays were carried out using the Amplex Red Hydrogen Peroxide Assay Kit (Molecular Probes, Inc.). This enzyme-coupled assay uses the fluorogenic Amplex Red reagent (10-acetyl-10*H*-phenoxazine-3,7-diol) in combination with horseradish peroxidase to detect trace amounts of hydrogen peroxide.[20] To determine the stoichiometry of H_2O_2 production per GFP chromophore generated, an endpoint analysis was carried out by absorbance. Urea-denatured

(25) Weast, R. C., Ed. *Handbook of Chemistry and Physics*; CRC Press: Cleveland, OH, 1974.
(26) Weiss, R. F. *Deep-Sea Res.* **1970**, *17*, 721−735.
(27) Xu, H.; Aylott, J. W.; Kopelman, R.; Miller, T. J.; Philbert, M. A. *Anal. Chem.* **2001**, *73*, 4124−4133.
(28) Jung, S.-K.; Gorski, W.; Aspinwall, C. A.; Kauri, L. M.; Kennedy, R. T. *Anal. Chem.* **1999**, *71*, 3642−3649.
(29) Pettersen, E. O.; Larsen, L. H.; Ramsing, N. B.; Ebbesen, P. *Cell Proliferation* **2005**, *38*, 257−267.
(30) Wachter, R. M.; Brett, A. K.; Heim, R.; Kallio, K.; Tsien, R. Y.; Boxer, S. G.; Remington, S. J. *Biochemistry* **1997**, *36*, 9759−9765.

GFP-trix was induced to mature by the rapid dilution technique described above, using a dilution factor of 22. Reducing agents such as TCEP were not added since they interfere with the assay. After 2 h, the maturation reaction was spun to remove aggregates, and 50 μL was added to 50 μL freshly prepared working solution containing the Amplex Red reagent and peroxidase, and the samples were incubated at room temperature in the dark for 30 min. The absorbance of samples and standards (0 to 1.0 μM H_2O_2) was read at 571 nm. All data were corrected for autoxidation of the reagent by subtracting the signal produced in the absence of any protein or H_2O_2, in reactions that were otherwise identically treated. To test whether additional H_2O_2 would be released from the protein upon denaturation, 22 μL of 0.5 M HCl was added to 90 μL of the maturation mix. The sample was subsequently neutralized by addition of 8 μL of 1 M NaOH before addition of the working solution. Chromophore concentration in the maturation mix after 2 h of stirring at 30 °C was determined by absorbance at 489 nm after spinning to remove aggregates, using an extinction coefficient of $\epsilon_{489} = 33\,787$ determined for GFP-trix in our laboratory. In a typical experiment, resorufin and chromophore concentrations ranged from 0.3 to 1.0 μM depending on the concentration of starting material.

Hydrogen peroxide production was also monitored in continuous mode using the fluorescence signal of resorufin (excitation 550 nm, emission at 580 nm), which is generated by the reaction of Amplex Red with hydrogen peroxide in the presence of HRP.[20] For this purpose, a 300 μL maturation reaction was set up in the fluorimeter cuvette with inclusion of Amplex Red/HRP working solution. A 240 μL folding buffer (50 mM HEPES pH 7.9, 300 mM NaCl, 1 mM EDTA) was mixed with a 50 μL Amplex Ultra working solution (90 μM Amplex Red, 0.12 units/mL HRP), and 10 μL urea-denatured immature GFP was added to the cuvette while stirring. The fluorimeter integration time per data point was set to 0.5 s. For the first 10 min, a data point was collected every 20 s, then every 30 s up to 1 h, then every 60 s up to 2 h. The fluorimeter slit width was set to 1 nm for both excitation (550 nm) and emission (580 nm). Fluorescence data were also collected for hydrogen peroxide standards under the same conditions for 2 h, at 0.0, 0.1, 0.2, and 0.3 μM H_2O_2. The reaction progress was approximated by pseudo-first-order reaction kinetics, since both Amplex Red and HRP were present in excess amounts. The data were computer-fitted to the equation $[F = F_{max} - e^{(-kt)} \cdot F_{max}]$, where F = fluorescence emission intensity, using the program Kaleidagraph. The average and sample standard deviation of the three rate constants for H_2O_2 production was determined (termed $k_4 \pm s$) and used in subsequent curve-fitting procedures (see below).

Maturation Kinetics Monitored by Chromophore Fluorescence. Inclusion body-derived protein was solubilized in urea and affinity-purified as above. A 300 μL aliquot of folding buffer (50 mM HEPES pH 7.9, 300 mM NaCl, 1 mM EDTA) was transferred to a 500 μL four-sided quartz cuvette placed into the thermostated fluorimeter sample holder. The buffer was stirred rapidly with a 3 mm stir bar and equilibrated to 30 °C with a circulating water bath. Protein maturation was initiated by addition of a 10 μL urea-denatured protein. All fluorescence experiments were carried out on a Jobin Ivon Horiba FluoroMax-3 fluorimeter. The excitation wavelength was set to 480 nm, and emission was monitored at 510 nm, with excitation and emission band pass set to 0.84 and 0.42 nm, respectively. The integration time per data point was 0.5 s. For the first 10 min after initiation of folding, a data point was collected every 20 s, thereafter every 30 s up to 60 min, then every 60 s up to 90 min to avoid photobleaching.

Global Curve Fitting and Error Estimate. Reaction progress curves for resorufin and GFP chromophore production were fitted to a minimal mechanism for GFP maturation using the program DynaFit.[21] The previously determined rate constant for peroxide production, k_4, was fixed to its average value, and the variable parameters k_1, k_2, and k_3 were extracted from curve fitting procedures. The fitting procedure was repeated with $k_4 + s$ and $k_4 - s$ (s = standard deviation), and the larger of the change in each fitted parameter k_1, k_2, and k_3 was used to estimate the error on that parameter.

Maturation Kinetics Monitored by HPLC. Starting material for in vitro maturation studies consisted of inclusion body-derived GFP-trix that was purified by Ni affinity immediately before carrying out the kinetic study. The denatured, immature protein was solubilized in 8 M urea, 50 mM HEPES pH 7.9, 50 mM NaCl, and 100 mM imidazole. Folding buffer consisted of 50 mM HEPES pH 7.9, 300 mM NaCl, and 1 mM EDTA. Several clear, four-sided 5 mL cuvettes were equipped with Teflon-coated flea stir bars, and 860 μL folding buffer and 100 μL 10 mM TCEP (final concentration 1 mM) were added to each, followed by equilibration at 30 °C using a Fisher Isotemp circulating waterbath and a tight-fitting cuvette water jacket. Protein maturation was induced by addition of 40 μL denatured immature protein to each cuvette containing 960 μL of rapidly stirring folding solution (25-fold dilution of urea-solubilized protein, about 10 μg/mL protein concentration in the maturation reaction).

At desired timepoints, the maturation reaction was filtered through 0.1 μm Millipore spin filter devices, followed by passage over a Sephadex G-75 column (300 μL slurry in a Pasteur pipet). For each maturation time point, a total volume of 5 mL of maturation mix was injected into a reverse-phase C4 semipreparative HPLC column (100 $\times$ 250 mm, Higgins Analytical, Inc.) to ensure sufficient material for subsequent trypsinolysis. The protein was eluted using a linear water-acetonitrile gradient containing 0.1% TFA. Protein fractions were hand-collected and lyophilized. Control injections consisted of immature urea-denatured GFP-trix, mature GFP-trix in the native state, and mature GFP-trix after urea and heat denaturation. Native GFP-trix of known concentration was used as standard. The absorbance was monitored at 280 nm and quantitated by peak area integration and subsequently was used to determine protein content in inclusion body-derived pools. Native GFP-trix was also used to determine the OD380/OD280 ratio of mature protein to be 1.38 at the apex of the eluting protein peak, as determined by the diode array optics of our HPLC system.

Trypsinolysis, Peptide Purification, and MALDI. Lyophilized fractions were digested with trypsin (TCPK-treated bovine, Sigma) essentially as described,[13] and tryptic peptides were separated by reverse-phase HPLC on a C18 analytical column (Vydac). Fractions were hand-collected, lyophilized, and subjected to MALDI-MS . Data were obtained using a Voyager DE STR mass spectrometer, and all instrumental settings and procedures were exactly as described previously.[13] Calibration was performed using CalMix 2 (Applied Biosystems) as an external standard. Initially, masses were determined for all peptide fractions eluting off the column and were assigned to predicted GFP-trix tryptic peptides. Since all peptides containing the chromophore-forming residues 65−67 eluted over a specific range of retention times, only this range was examined by MALDI during later runs.

Acknowledgment. This work was supported by the National Science Foundation (NSF), Grant MCB-0213091 (to R.M.W.). NSF Grant CHE-0131222 provided funds to purchase the mass spectrometer.

Supporting Information Available: Representative MALDI mass spectra. Experimental description of tryptophan fluorescence experiments. Tryptophan fluorescence spectra as a function of maturation time. This material is available free of charge via the Internet at http://pubs.acs.org.

JA0580439