BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

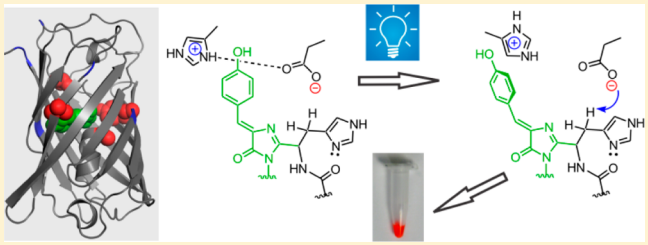
Acid–Base Catalysis and Crystal Structures of a Least Evolved Ancestral GFP-like Protein Undergoing Green-to-Red Photoconversion

Hanseong Kim,[†] Timothy J. Grunkemeyer,[†] Chintan Modi,[‡] Liqing Chen,[†] Raimund Fromme,[†] Mikhail V. Matz,[‡] and Rebekka M. Wachter*[,†]

[†]Department of Chemistry and Biochemistry, Arizona State University, Tempe, Arizona 85287, United States

[‡]Section of Integrative Biology, University of Texas, Austin, Texas 78712, United States

ABSTRACT: In green-to-red photoconvertible fluorescent proteins, a three-ring chromophore is generated by the light-activated incorporation of a histidine residue into the conjugated π-system. We have determined the pH–rate profile and high- and low-pH X-ray structures of a least evolved ancestor (LEA) protein constructed in the laboratory based on statistical sequence analysis. LEA incorporates the minimal number of substitutions necessary and sufficient for facile color conversion and exhibits a maximal photoconversion quantum yield of 0.0015 at pH 6.1. The rate measurements provide a bell-shaped curve, indicating that the reaction is controlled by the two apparent pK_a values, 4.5 ± 0.2 and 7.5 ± 0.2, flanking the chromophore pK_a of 6.3 ± 0.1. These data demonstrate that the photoconversion rate of LEA is not proportional to the A-form of the GFP-like chromophore, as previously reported for Kaede-type proteins. We propose that the observed proton dissociation constants arise from the internal quadrupolar charge network consisting of Glu222, His203, Glu148, and Arg69. Increased active site flexibility may facilitate twisting of the chromophore upon photoexcitation, thereby disrupting the charge network and activating the Glu222 carboxylate for the abstraction of a proton from a carbon acid. Subsequently, the proton may be delivered to the Phe64 carbonyl by a hydrogen-bonded network involving Gln42 or by means of His65 side chain rotations promoted by protein breathing motions. A structural comparison of LEA with the nonphotoconvertible LEA-Q42A variant supports a role for Gln42 either in catalysis or in the coplanar preorganization of the green chromophore with the His65 imidazole ring.



Green-to-red (G/R) photoconvertible proteins constitute a class of GFP-like proteins that are able to undergo light-dependent color conversion. The green form (emission at ~518 nm) is irreversibly modified to a red-fluorescing form (emission at ~582 nm) when exposed to UV or violet radiation, as first described for Kaede.[1] Photoconvertible relatives of Kaede include EosFP,[2] DendFP,[3,4] mcavRFP and rfloRFP,[3,5] and others. Because of its color switching properties, this class of fluorescent proteins (FPs) has become an important tool in the development of super-resolution microscopic techniques for imaging subcellular biological structures with unprecedented detail.[6] The field of super-resolution microscopy has expanded dramatically in recent years and continues to fuel an intense and ever-growing interest in genetically encodable photochromic fluorescent probes. Regardless, unraveling the catalytic mechanism of light-triggered G/R color conversion has remained challenging, in spite of a significant effort put forth by a number of research groups. Although several recent review articles have provided good summaries of prevalent mechanistic hypotheses,[7,8] progress toward a more complete understanding of the reaction progress has been slow, in large part because of the difficulty of generating sufficiently large populations of intermediate states that can be trapped or monitored in some way. Here, we present a novel approach that involves ancestral gene reconstruction technology to better understand how the protein fold catalyzes the light-induced bond rearrangements necessary to yield a red-emitting chromophore.

On the basis of phylogenetic analyses, three clades of GFP-like proteins (B, C, and D) have been distinguished in stony corals (class Anthozoa, order Scleractinia).[9] In clade D, extant cyan, green, and red fluorescent proteins (FPs) are well-represented; however, red fluorescent proteins in this clade are solely of the green-to-red (G/R) photoconvertible type (Kaede type).[10] Therefore, the great star coral *Monastraea cavernosa* (clade D suborder Faviina) was originally chosen as a model organism for studying the evolution of green-to-red color transitions.[10] A combinatorial approach was employed to synthesize the set of sequences consistent with an ancestral node inferred to be the common ancestor of all color classes expressed in this organism.[10] Because all reconstructed sequences encoded proteins with green (505 nm) emission,[9,10] the ancestral protein was termed ALL-GFP. A sequence

Received: July 25, 2013
Revised: October 16, 2013
Published: October 17, 2013

dx.doi.org/10.1021/bi401000e | *Biochemistry* 2013, 52, 8048–8059

comparison of extant red fluorescent proteins of *M. cavernosa* with the green ancestral node demonstrated that the least divergent pair involved a total of 37 residue substitutions. Therefore, a large number of transitional libraries were constructed, with each of the 37 sites given an ∼50% chance of bearing the ancestral residue and an ∼50% chance of bearing the extant residue.[11] A statistical analysis allowed for the identification of 13 sites (including one deletion) that appeared to play a critical role in the evolutionary transition from the green to the red phenotype. The substitutions both necessary and sufficient to endow the ancestral ALL-GFP with efficient G/R photoconversion properties consisted of A63V, Q65H, T72A, T109R, S110N, Y121N, M162T, V165I, R204C, R227H, delY[227−228], and M228G (residue numbers based on the conventional GFP numbering system).[11] The evolved ancestor bearing this set of residue replacements was termed the least evolved ancestor (LEA). Although Q65H is absolutely required for the generation of red fluorescence, because the histidine becomes incorporated into the chromophore upon photoconversion, this substitution alone results in retention of the green phenotype. Among the set of essential residues, five sites concern residues with side chain positions that cluster near the chromophore: A63V, Q65H, S110N, Y121N, and V165I (Figure 1, red residues);[11] the remaining eight sites concern residues with side chains that either point toward the exterior solvent or are located along the C-terminal tail (Figure 1, blue residues).

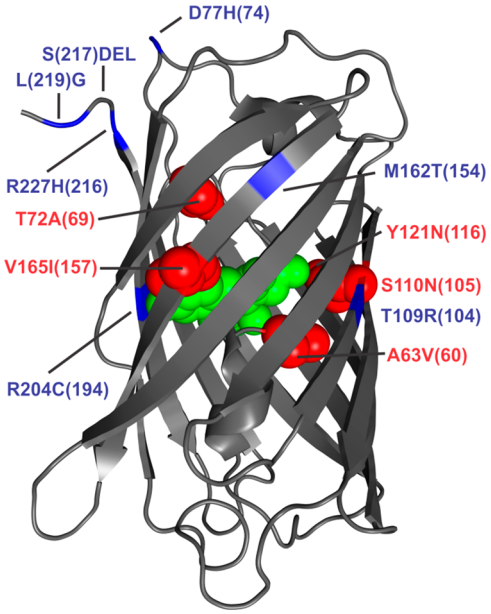


Figure 1. Protein fold of one chain of the LEA tetramer. The residue positions are labeled according to the GFP consensus sequence, with the LEA residue number given in parentheses. The 11 residue substitutions that correlate with photoconversion are highlighted, and the location of the ALL-GFP Ser217 deletion is indicated. Red denotes side chains clustering around the chromophore and blue side chains either solvent-exposed or part of the subunit interface.

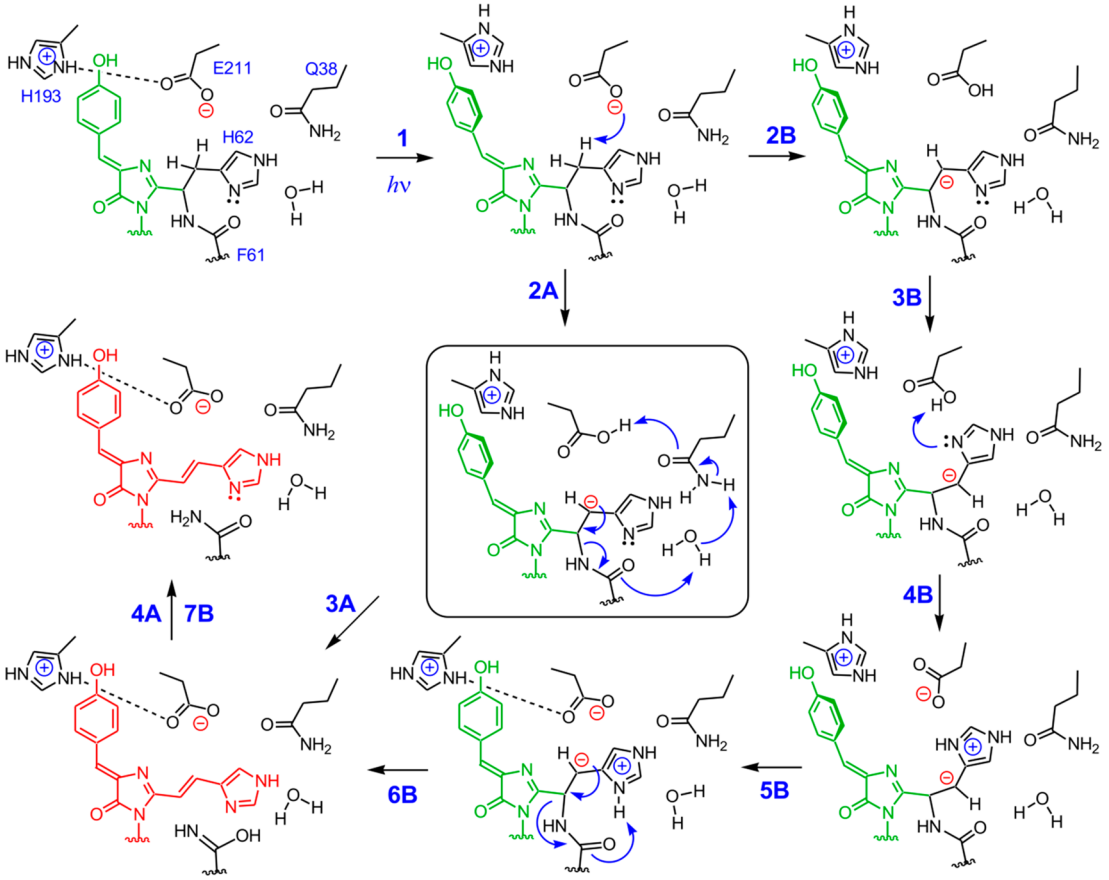
Although the mechanism of photoconversion in Kaede-type proteins continues to be heavily debated, several features have been shown to be common to all members of this class. Kaede-type proteins always bear a His-Tyr-Gly tripeptide in the interior of the 11-stranded β-barrel. The three-ring red chromophore is constructed from the GFP-like chromophore by the introduction of π-orbital overlap with the His65 imidazole ring (His62 according to LEA and Kaede numbering).[12] The conjugation of the aromatic ring systems is accomplished by desaturation of the His65 $C_\alpha-C_\beta$ bond, coupled to a main chain bond scission adjacent to the chromophore (Scheme 1). In the β-elimination reaction, the His65 amide nitrogen is thought to be ejected from the α-carbon of the same residue. Although the interior backbone cleavage results in N-terminal 10 kDa and C-terminal 18 kDa peptide fragments, the protein remains natively folded. The reaction has long been thought to occur only by direct excitation of the protonated A-form of the GFP-like chromophore.[2] On the basis of this idea, an excited-state proton transfer (ESPT) has been proposed to occur from the chromophore's phenolic hydroxyl to the imidazole ring of His65 over a distance of >10 Å.[2,12] According to this model, the ensuing imidazolium cation would serve as a proton donor to the amide leaving group in the β-elimination reaction. Although this model is attractive, experimental support for long-range ESPT has not been forthcoming, and a hydrogen bond network connecting the chromophore's phenolic group with His65 has not been identified in any X-ray structures.[13]

A two-step β-elimination mechanism has been proposed, as supported by X-ray structures of green and red forms of an engineered protein termed KikGRX.[12,14] A structural comparison of these forms has suggested that the histidine *cis* isomer may be generated by a $C_\alpha-C_\beta$ bond rotation of a reaction intermediate.[15] Although an alternate, one-step concerted mechanism has also been considered,[13] several more recent observations have provided arguments against this type of mechanism.[15−17] The two-step model has recently been expanded by means of quantum mechanics/molecular mechanics computation of the reaction progress, and two sequential mechanisms have been proposed to compete with each other.[18] In the first, main chain cleavage would be followed by abstraction of a proton from C_β of His65, and in the second, proton abstraction would be followed by main chain cleavage.[18] In each of these mechanisms, C−N bond scission is thought to be facilitated by the transfer of a proton from the imidazolium cation of His65 to the amide carbonyl of Phe64. In support of this notion, the proximity of Nδ1 of His65 to the carbonyl O of Phe64 is consistent with direct hydrogen bonding,[13] albeit with poor angular geometry.[7] However, to date, the thermodynamic driving force for His65 protonation has remained unclear. Although computer calculations have led to the suggestion that a neutral imidazole could also serve as a proton donor to yield a transient imidazolide anion within the lifetime of the chromophore's excited state,[17] it appears that the chemical steps would more likely proceed after the system has decayed back to the electronic ground state.[18]

To improve our understanding of the mechanism of G/R photoconversion in Kaede-type proteins, we have measured photoconversion rates of the reconstructed LEA protein as a function of pH. In addition, we have determined several high-resolution X-ray structures of LEA. On the basis of these data, we propose a novel mechanistic model that involves light-triggered chromophore twisting to disrupt nearby charge−charge interactions, thus enhancing the basicity of Glu222. Subsequently, the proton may be transferred to Phe64 via a proton hopping or proton shuttling mechanism (Scheme 1).

dx.doi.org/10.1021/bi401000e | *Biochemistry* 2013, 52, 8048−8059

Scheme 1. Proposed Mechanisms for the Photoconversion of Kaede-Type Fluorescent Proteins, (A) Proton Relay by Means of Gln42 and (B) His65 as a Proton Shuttle



EXPERIMENTAL PROCEDURES

Protein Expression and Purification. The gene encoding the reconstructed least evolved ancestor protein (LEA) was inserted into expression plasmid pGEM-T (Promega, Inc.), C-terminally tagged with six His residues, and expressed in *Escherichia coli* BL21(DE3). At an OD_{600} of 0.6, the temperature was reduced from 37 to 25 °C, the culture was induced with IPTG, and cells were allowed to grow overnight. Proteins were purified by Ni-NTA (Qiagen) affinity chromatography, immediately dialyzed against buffer containing 50 mM HEPES (pH 7.9), 20 mM NaCl, and 1 mM EDTA, and flash-frozen. During expression, purification, and storage, exposure to room light was minimized by wrapping all equipment, beakers, columns, and trays in aluminum foil. Floor shaker lid windows and glass refrigerator doors were obscured as well.

Crystallization and Data Collection. Crystals of LEA were grown by hanging drop vapor diffusion at 20 °C in the dark. The drops consisted of 2 μL of protein [10 mg/mL in 50 mM HEPES (pH 7.9), 20 mM NaCl, and 1 mM EDTA] and 2 μL of mother liquor. For LEA crystallized at high pH, the mother liquor consisted of 70 mM Tris (pH 7.6), 100 mM sodium acetate, 12% PEG 4000, and 15% glycerol. For LEA crystallized at low pH, the mother liquor consisted of 120 mM glycine-phosphate buffer (pH 3.8), 100 mM magnesium chloride, and 20% PEG 4000. For LEA-Q38A, the mother liquor consisted of 85 mM Tris (pH 7.5), 85 mM sodium acetate, and 16% PEG 4000. Crystals were flash-frozen in liquid nitrogen, and all X-ray diffraction data were collected at 100 K. The LEA-Q38A data set was collected using an R-AXIS IV^{++} image plate detector mounted on a Rigaku RU200HB rotating anode generator equipped with osmic confocal mirrors. Data sets for the high-pH and the low-pH forms of LEA were collected at the Advanced Photon Source (Argonne National Laboratory, Argonne, IL) beamline ID-19, using an ADSC Q315R detector.

Structure Determination and Refinement. *High-pH Form of LEA.* The data set was processed using the HKL2000 software package.[19] Molecular replacement was conducted with PHASER,[20] using the X-ray structure of the A-chain of EosFP [Protein Data Bank (PDB) entry 1ZUX][13] as a search model. The initial electron density was autotraced with BUCCANEER[21] to introduce the correct LEA amino acid sequence. Side chain positions were adjusted manually within COOT.[22] After 5% of the data had been removed for R_{free} calculations, rigid body refinement was conducted using REFMAC.[23] Positional refinement was first conducted in the absence of a chromophore model. After several rounds of refinement to the full resolution limit, the chromophore was modeled by using the chromophore coordinates from the cmFP512 X-ray structure (PDB entry 2C9J)[24] edited to reflect the Q65H substitution. The chromophore was subsequently refined against an appropriate target geometry file as described previously.[25] At this point, isotropic *B* factor refinement was conducted prior to water modeling. Subsequently, solvent molecules were modeled into strong electron density ($2F_O - F_C$ at the $\geq 1\sigma$ level) within hydrogen bonding distance of protein groups. Final rounds of refinement were conducted in

dx.doi.org/10.1021/bi401000e | *Biochemistry* 2013, 52, 8048−8059

REFMAC with inclusion of solvent molecules and with anisotropic *B* factor refinement.

Low-pH Form of LEA. To better compare the low-pH structure with the high-pH structure, the initial procedures used to process the low-pH data set were chosen to be identical to those used for the high-pH data set. The data were processed using the HKL2000 software package,[19] and molecular replacement was conducted with PHASER,[20] using the X-ray structure of the A-chain of EosFP (PDB entry 1ZUX)[13] as a search model. Initial rigid body refinement was followed by several rounds of NCS-restrained refinement of the tetramer in the asymmetric unit, which was followed by TLS (translation−libration−screw motion) refinement without modeling any solvent molecules or chromophores. For TLS refinement, TLS parameters were produced by TLSMD,[26] each protein chain was defined as one TLS group, and *B* factors were initially set to 40. One round of TLS refinement consisted of 20 cycles, which was followed by 10 cycles of maximum likelihood restrained refinement in REFMAC, to adjust atomic coordinates and anisotropic *B* factors while keeping the TLS parameters constant. After several rounds, water molecules and chromophores were added to the model, and additional restrained refinement was performed. Because the map of chain D continued to be of rather poor quality, only chains A−C were used for structural analysis.

LEA-Q38A. The data set was processed using MOSFLM[27] and scaled and merged using SCALA in the CCP4 suite of programs.[28] Molecular replacement was conducted with PHASER,[20] using the X-ray structure of LEA as a search model. After the initial refinement, Gln38 was replaced with Ala using COOT.[22] The chromophore was modeled by use of the LEA chromophore coordinates. All other computational procedures were conducted as described for the high-pH form of LEA.

Measurement of the Absorbance, Fluorescence, and Chromophore pK_a. Absorbance spectra were recorded on a Shimadzu UV-2401 spectrophotometer between pH 4.0 and 13.0 at room temperature. Protein stock solutions (1 mg/mL) were diluted 10-fold into appropriate buffers and allowed to equilibrate. pH 4.0−11.0 buffers consisted of 50 mM citrate, acetate, PIPES, HEPES, CHES, or CAPS, all containing 100 mM NaCl and 1 mM EDTA. Buffers at pH 12 and 13 were prepared by adjusting NaOH concentrations to 10 and 100 mM. The chromophore pK_a of each variant was determined by curve fitting the absorbance intensity of the chromophore anion [band B (Table 2)] to the Henderson−Hasselbalch equation using Kaleidagraph. The data were fit to eq 1

$$A = A_{max}/[1 + 10^{(pK_a - pH)n}] \quad (1)$$

where A_{max} represents the absorbance upon full ionization, K_a represents the apparent proton dissociation constant for the titrating site(s), and n represents the Hill coefficient (n_H).

Fluorescence excitation and emission spectra were recorded using a SPEX Fluoromax-3 (Jobin Yvon Horiba) fluorimeter with the slit width set to 1.0 nm and an integration time of 1 s. The samples contained 0.1 mg/mL protein in 50 mM HEPES (pH 7.9), 20 mM NaCl, and 1 mM EDTA buffer.

Determination of the Quantum Yield of Green Fluorescence. The quantum yield of green fluorescence (Φ_{Fl}) of protein solutions that had been kept in the dark was determined using fluorescein as a standard (1 μM in 0.1 M NaOH). The buffer consisted of 50 mM HEPES (pH 7.0), 50 mM NaCl, and 1 mM EDTA for all protein preparations. A new fluorescein standard was prepared for each quantum yield determination. The crossover point of protein and fluorescein absorbance spectra was utilized as the excitation wavelength for fluorescence emission scans recorded for the same sample (integration time of 1.0 s, increment of 1.0 nm, and slit width of 1.0 nm). The emission scans were integrated with the left integration limit set to the wavelength at which one-third of the maximal fluorescence intensity was observed. The integration results were fit to eq 2

$$\Phi_{Fl} = \Phi_R \frac{I}{I_R} \times \frac{OD_R}{OD} \quad (2)$$

where Φ_{Fl} is the quantum yield, I is the integrated intensity, and OD is the optical density at the respective excitation wavelength. The subscript R stands for the reference fluorophore with a known quantum yield of 0.95.

Measurement of Photoconversion Kinetics. Protein samples (0.1 mg/mL) were prepared by combining 100 μL of a 1 mg/mL protein stock solution (kept in the dark) and 900 μL of buffer. Samples were incubated overnight in the dark. To examine the pH dependence of LEA photoconvertion, proteins were illuminated with 405 nm light at 13 different pH values for a total of 60 min. The illumination apparatus consisted of three tight-tolerance UV LED lamps (BIVAR) emitting 405 ± 2.5 nm light with a 10 mW emitting power at a 30° viewing angle. These lamps were mounted immediately adjacent to each other in a home-built dark box equipped with a 25 W power supply and an on−off switch. One milliliter of a 0.10 mg/mL protein solution was transferred into a 3 mL disposable UV cuvette, which was placed into the dark box such that the LEDs were positioned directly adjacent to the cuvette window. Therefore, the path length of light passing through the solution was 1 cm. During the reaction, absorbance spectra from 250 to 650 nm were collected at specific time points (0, 1, 2, 3, 5, 7, 10, 15, 20, 30, 45, and 60 min) using a Shimadzu UV-2401 spectrophotometer. The progression of the photoconversion reaction was determined by monitoring the absorbance loss of the A- and B-bands of the GFP-like chromophore. The absorbance values at 386 nm (A) and 504 nm (B) were utilized to calculate the concentration of the total GFP-like chromophore using their respective extinction coefficients (A-band ε = 22943 M^{-1} cm^{-1}; B-band ε = 86754 M^{-1} cm^{-1}). The fractional loss of the GFP-like chromophore was plotted as a function of time and computer-fitted to first-order reaction kinetics using Kaleidagraph (Synergy). Because a single-exponential model did not provide acceptable fits to the data, a double-exponential fit was conducted according to eq 3.

$$A = A_1 e^{-k_1 t} + (1 - A_1) e^{-k_2 t} \quad (3)$$

where A is the fraction of GFP-like chromophore remaining, A_1 is the fractional chromophore population decaying with a fast rate constant k_1, and $1 - A_1$ is the fractional population decaying with a slow rate constant k_2.

The observed fast and slow rate constants (k_1 and k_2, respectively) were plotted as a function of pH. Because k_1 exhibited a pH−rate profile different from that of k_2, different equations were employed for curve fitting. The values of k_1 were fit using eq 4, which describes an acid-catalyzed monoprotic system, in which the observed rate constant depends on an intrinsic (pH-independent) rate constant k_{int}

and a protein titratable group with an apparent acid dissociation constant K_a.

$$k_1 = k_{int}\frac{10^{-pH}}{K_a + 10^{-pH}} \quad (4)$$

The values of k_2 were fit using eq 5. Equation 5 describes a simple diprotic system, in which the two successive acid dissociation constants, K_{a1} and K_{a2}, are within 3.5 pH units of each other.[29] As before, k_{int} represents the intrinsic (pH-independent) rate constant.

$$k_2 = \frac{k_{int}K_{a1}[H^+]}{K_{a1}[H^+] + [H^+]^2 + K_{a1}K_{a2}} \quad (5)$$

Determination of the Photoconversion Quantum Yield. Because the UV-LED lamps used for photoconversion emit 405 nm light, the 405 nm absorbance was measured for a sample of LEA protein (0.1 mg/mL, pH 4.03) that had not previously been exposed to light. On the basis of its chromophore pK_a value (Table 3), 99.5% of the chromophore exists in the protonated form (A-band) at this pH value. The same sample was subsequently used for photoconversion kinetic experiments. The fraction of light absorbed at 405 nm (band A) was calculated from the absorbance value. The LED-illuminated surface area of the protein solution was determined to be 0.86×10^{-4} m^2. A light meter (LiCor) was employed to determine the number of photons per second reaching the solution surface. From these values, the number of photons absorbed per second was calculated. The concentration of the protonated chromophore (band A) was determined to be 3.21 μM on the basis of its 382 nm absorbance. Subsequently, rate constant k_2 for the photoconversion of LEA at pH 4.03 was utilized to calculate the number of chromophores converted per second in a volume of 1 mL. The ratio of chromophores converted to photons absorbed provided the LEA quantum yield of G/R photoconversion at pH 4.03 ($QY_{pc} = 2.45 \times 10^{-4}$). A similar procedure was used to calculate the quantum yield at pH 6.15 ($QY_{pc} = 1.50 \times 10^{-3}$).

Determination of the Protein Oligomerization State in Solution. The oligomeric state of LEA protein was assayed by size-exclusion high-performance liquid chromatography (HPLC), using multiangle light scattering and refractometry detection at 25 °C. Oligomeric species were separated on a G3000PW$_{XL}$ column [TosoHass, 300 mm × 7.8 mm (Tosoh Bioscience LLC, King of Prussia, PA)], using an injection volume of 20 μL and a flow rate of 0.4 mL/min. The molar mass was determined using an EOS photometer and an Optilab rEX differential refractometer (Wyatt Technology, Santa Barbra, CA). Data analysis was conducted using Astra version 4.9 (Wyatt Technology).

RESULTS

Biochemical Features of the Reconstructed Least Evolved Ancestor (LEA) Protein. To better understand the structural requirements for G/R photoconversion, we expressed and purified the reconstructed LEA protein in the dark. The quantum yield for green fluorescence was determined to be 0.81 (eq 2), and LEA was demonstrated to undergo efficient G/R photoconversion by illumination with LED-generated 405 nm light. Compared to its green ancestor, LEA bears five substituted sites clustered near the chromophore and eight substituted sites more remote from the chromophore (Figure 1).[11] In line with convention, GFP residue numbering is used throughout this work, with the actual residue number of LEA provided in parentheses. To better understand the relationship between chromophore charge state and photochemistry, spectrophotometry and curve fitting (eq 1) were used to determine the pK_a value of the green LEA chromophore. This value was shown to be equal to 6.3 ± 0.1, similar to that reported for Kaede (pK_a = 5.6)[1] and EosFP (pK_a = 5.8)[2] (Table 1). However, all three LEA variants prepared for the analysis of mechanistic features [LEA-Q42(38)A, LEA-Q42(38)M, and LEA-Q42(38)N] displayed significantly higher pK_a values (7.8, 7.8, and 7.5, respectively), similar to those reported for Dendra2 (pK_a = 7.1)[30] and KikGRX (pK_a = 8.0).[15]

Table 1. Acid−Base Equilibria and Approximate Crystallization pH of LEA and Its Variants[a]

	GFP-like chromophore pK_a ± standard deviation	Hill coefficient (n_H)	crystallization pH
LEA	6.3 ± 0.1	0.60	7.6 (high pH) ~6.4 (low pH)
LEA-Q42A	7.8 ± 0.1	0.62	7.5
LEA-Q42M	7.8 ± 0.1	0.67	–
LEA-Q42N	7.5 ± 0.1	0.76	–

[a]The chromophore pK_a value and Hill coefficient were determined using the Henderson−Hasselbalch equation, and the quantum yield was determined using fluorescein as a standard. All measurements were made on the "green", pre-photoconversion state of the protein.

pH-Dependent Photoconversion Kinetics and Photoconversion Quantum Yield of LEA. To examine the pH dependence of photoconversion, solutions of LEA were illuminated with LED-generated 405 nm light for a total duration of 60 min. At this wavelength, only the A-band of the GFP-like chromophore is electronically excited. The 386 nm (A-band, neutral) and 504 nm (B-band, anionic) peaks were converted to new peaks absorbing at ~460 nm (neutral red) and 571 nm (anionic red) in a time- and pH-dependent manner (Figure 2). For each pH value, the fractional loss of the GFP-like chromophore (A- and B-band populations) was plotted as a function of exposure time (Figure 3A,B). This procedure was chosen because monitoring the increase in the magnitude of the 460 nm band (neutral red) was not feasible due to substantial peak overlap with the green B-band, and monitoring the increase in the magnitude of the 571 nm band (anionic red) was feasible only at the highest pH values tested. Curve fitting to a double-exponential equation based on first-order kinetic models (eq 3) provided low residuals and allowed for the extraction of fast and slow apparent rate constants (k_1 and k_2, respectively) at each pH value. The fast process occurred within the first minute of light exposure and affected the A-band only, with ~9% loss of A-band absorbance at the lowest pH values in the absence of conversion to the B-band. Not surprisingly, fitting of k_1 to an acid-catalyzed reaction model (eq 4) demonstrated that this process slows with a midpoint equal to the chromophore pK_a value (data not shown), consistent with photobleaching of the A-band by 405 nm light. However, the actual values of k_1 are unreliable as they are primarily determined by two data points (Figure 3A). The loss of A-band intensity may be due to reversible photobleaching upon chromophore twisting, irreversible photobleaching via photodamage, or both.

dx.doi.org/10.1021/bi401000e | Biochemistry 2013, 52, 8048−8059

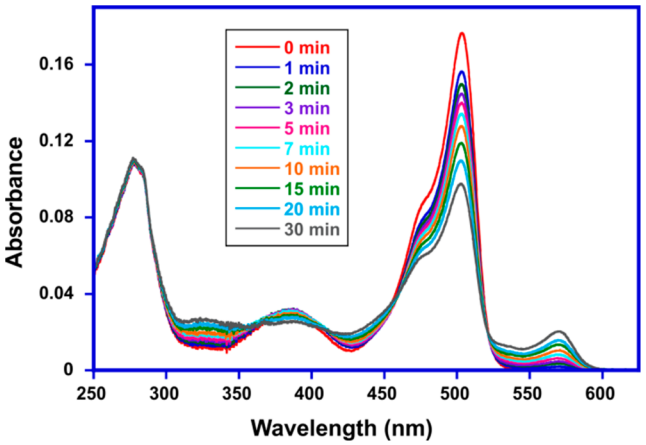


Figure 2. LEA absorbance spectra collected as a function of irradiation time using 405 nm LED light. The protein was kept in the dark and was equilibrated in pH 6.45 buffer overnight. Prior to irradiation, chromophore bands were observed at 386 nm (A-band) and 504 nm (B-band). Upon irradiation, the intensity of these bands decreased, and new bands were generated with maxima at ~460 nm (neutral red chromophore) and 571 nm (anionic red chromophore). In addition, new features were observed at ~330 nm.

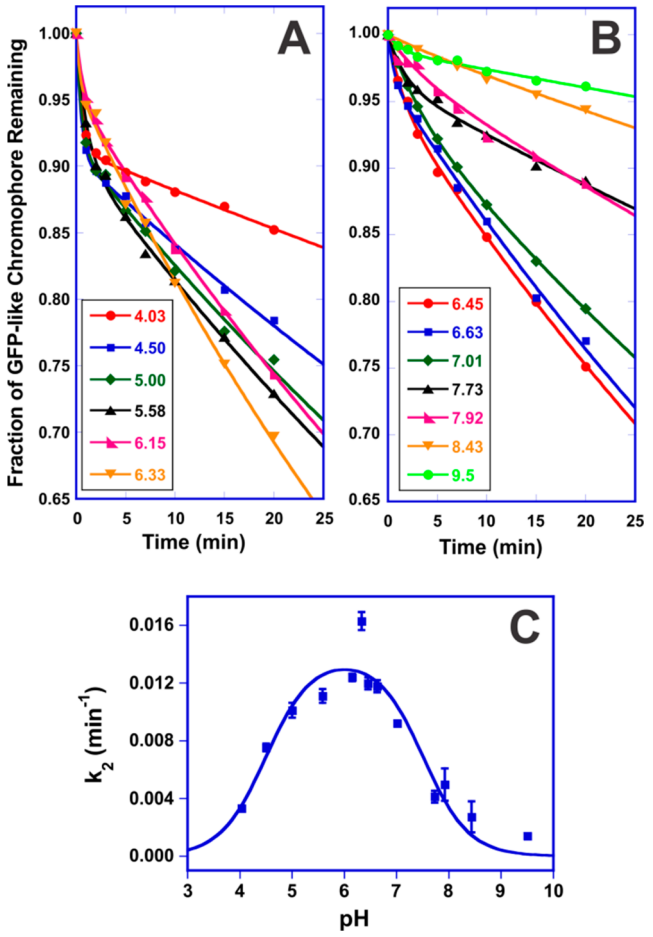


Figure 3. Fractional loss of the green chromophore of LEA upon 405 nm light irradiation as a function of time and pH. (A) Kinetic data between pH 4.0 and 6.3 and (B) kinetic data between pH 6.4 and 9.5. The measurement values are indicated by the symbols (see the legend) and the curve fits to eq 3 by the solid lines. (C) Values of k_2 and their error estimates, as extracted from the curve fits to eq 3, plotted as a function of pH (squares). The standard deviation for each data point is indicated by vertical bars. The k_2 values were fit to eq 5 to extract apparent acid dissociation constants pK_{a1} and pK_{a2} (solid line).

However, the k_2 values extracted by curve fitting to eq 3 (Figure 3A,B) correlated well with photoconversion to the red chromophore. At high pH, the kinetic constant determined by monitoring the loss of the green form was found to be similar to the value determined by monitoring the rise of the red anion (λ_{max} = 570 nm), supporting the notion that k_2 accurately reflects G/R photoconversion. Graphing of k_2 values as a function of pH provided a bell-shaped curve, providing evidence that both acid and base catalysis play a role in photochromism (Figure 3C). Therefore, the apparent rate constants were computer-fitted to an equation describing a "diprotic" system with pK_a values within 3.5 pH units of each other.[29] Using eq 5, the apparent pK_{a1} value was extracted to be 4.5 ± 0.2, the apparent pK_{a2} value to be 7.5 ± 0.2, and the intrinsic (pH-independent) rate constant (k_{int}) to be (2.3 ± 0.2) × 10^{-4} s^{-1}. The curve fit indicates that in LEA, the photoconversion rate reaches a maximum at pH 6.0 (Figure 3C). Using a light meter, the quantum yield of photoconversion (number of chromophores converted per photon absorbed) was calculated to be 1.5 × 10^{-3} at pH 6.2 and 2.4 × 10^{-4} at pH 4.0. These values are consistent with previously published values for Kaede-type proteins such as KikGR.[14,15]

Surprisingly, the apparent pK_{a2} value of 7.5 does not reproduce the pK_a value of 6.3 determined for the green chromophore of LEA (Table 1). Although several reports have suggested that photoconversion rates of Kaede-type proteins are directly proportional to the concentration of the A-band of the green form,[1,2,30] more recent work has questioned this notion with respect to some variants of EosFP and Dendra2.[31] In support of these observations, the pH−rate profile of LEA reported here suggests that the charge states of amino acid side chains, rather than the charge state of the chromophore itself, determine the observed rate. Therefore, we tentatively assign pK_{a1} to Glu222(211), which must be anionic to fulfill its role as a base, and pK_{a2} to His203(193), which must be protonated to maintain its ionic interactions. On the basis of the LEA crystal structure (see below), the His203(193) imidazolium cation forms ionic bonds with both Glu148(144) and Glu222(211) (Figure 4). Therefore, the pK_a of His203(193) is expected to be elevated, consistent with a value of 7.5, and linked to the pK_a values of the two glutamates. On the other hand, it would be difficult to argue that the apparent pK_a of 7.5 should be assigned to His65(62), because its imidazole group is located within a hydrophobic pocket lined by Met40 and Leu209. Electrostatic interactions that would stabilize the imidazolium cation are not observed in the LEA crystal structure. Therefore, the His65(62) acid dissociation constant is expected to be decreased below 6, rather than increased to 7.5.

Description of X-ray Structures of LEA. To avoid light-induced photoconversion, we expressed, purified, and crystallized all protein in the dark, so that the crystal structures represent the green pre-photoconversion state only. To examine potential structural differences as a function of chromophore charge state, the structure of LEA was determined to 1.8 Å at a crystallization reservoir pH of 7.6 (high pH) and to 1.5 Å resolution at a nominal reservoir pH of 3.8 (low pH). However, after appropriate dilution of an aliquot of the reservoir solution with the protein solution, the pH was measured to be 6.37 for the low-pH crystallization condition. Regardless, this value constitutes only an estimate of the crystal

dx.doi.org/10.1021/bi401000e | Biochemistry 2013, 52, 8048−8059

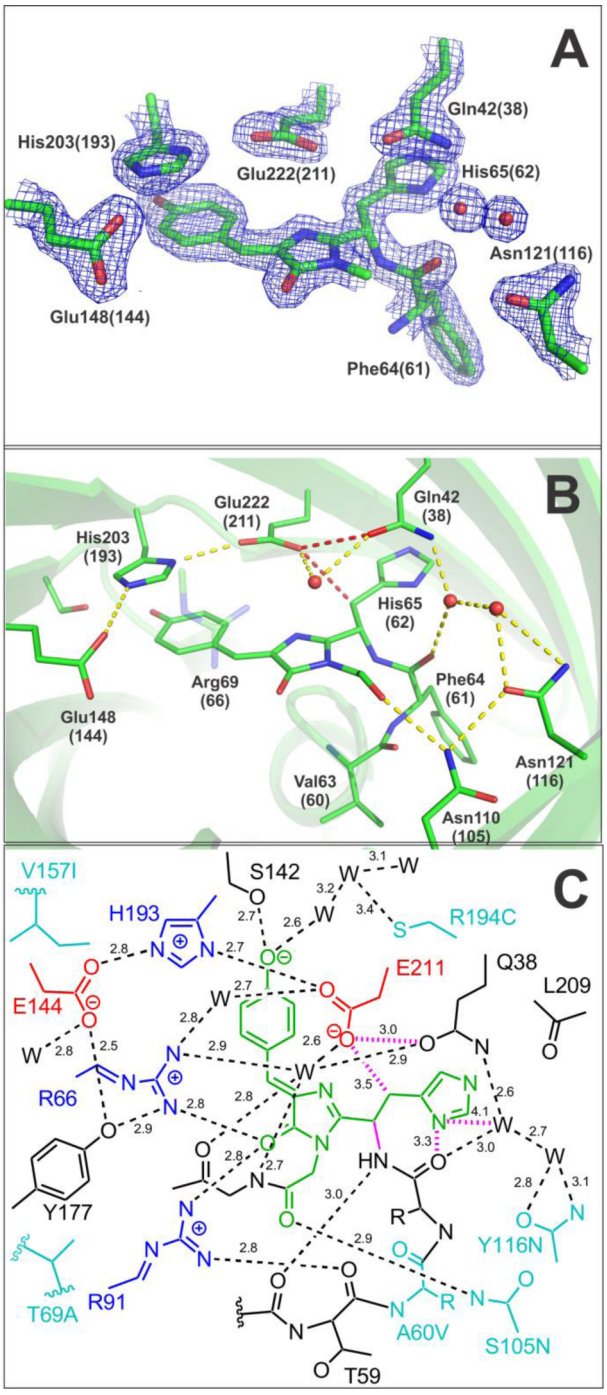


Figure 4. (A) Electron density map of the photoconvertible LEA variant. The $2F_O - F_C$ maps were contoured at 1.5σ. (B) LEA active site structure. Hydrogen bonds are shown as yellow dashed lines and close contacts as red dashed lines. Residue positions 121(116) and 110(105) correlate with photoconversion. The imidazole ring of His65(62) may be weakly H-bonded to the backbone carbonyl of Leu209 (not shown) but does not exhibit any other H-bonding interactions. (C) Schematic representation of the LEA chromophore and its environment. Black dashed lines indicate hydrogen bonds, and pink dashed lines indicate close contacts.

pH, as the mother liquor prepared for pH measurements was not equilibrated with crystals. In addition, recent work has demonstrated that under acidic conditions, the pH inside lysozyme crystals is between 0.3 and 1.0 pH unit lower than the bath pH.[32] As this phenomenon is thought to have more general relevance, the acidic LEA crystal may exhibit an internal pH value of <6. Unfortunately, crystals grown at lower pH values did not diffract, and crystals soaked into buffers below pH 6 dissolved.

For the high-pH LEA structure, the crystallization pH was ~1.3 units above the chromophore pK_a value (Table 1), thus representing the anionic B-state of the green chromophore. On the other hand, the low-pH LEA structure is likely composed of both A- and B-states. Although the reservoir solution pH is nearly equal to the LEA chromophore pK_a of 6.3, a lower pH value inside the crystal would favor the A-form. However, structural heterogeneity was not observed, and superposition of the high- and low-pH structures of LEA (α-carbon root-mean-square deviation of 0.40 Å) did not provide any significant differences in atomic positions. These data provide strong evidence that chromophore protonation does not induce side chain positional adjustments (data not shown). Therefore, the structure determined at pH 7.6 may be considered to be representative of both chromophore charge states.

High-pH LEA crystallized in space group $I222$ with one molecule per asymmetric unit, whereas low-pH LEA crystallized in space group $P2_12_12_1$ with four molecules per asymmetric unit (Tables 2 and 3). The relative protomer orientations in the tetramer generated by either crystallographic ($I222$) or noncrystallographic ($P2_12_12_1$) symmetry are almost identical to those found in EosFP and Kaede. As described previously for other Anthozoan GFP-like proteins,[33] the subunit interfaces consist of two types, the A−B interface with a nearly perpendicular arrangement and the A−D interface with a nearly antiparallel arrangement of β-barrels. The A−B interface is dominated by hydrophilic interactions, whereas the A−D interface is dominated by hydrophobic interactions. The C-terminal peptide known to frequently form a clasp around the neighboring subunit appears to be crystallographically disordered in LEA. Therefore, it is possible that the set of R204(194)C, R227(216)H, Y(217-DEL), and M(219)G

Table 2. Crystallographic Data Collection Statistics

	LEA (high pH)	LEA (low pH)	LEA-Q42A
space group	$I222$	$P2_12_12_1$	$I222$
no. of chains per asymmetric unit	1	4	1
unit cell			
a (Å)	48.77	82.44	49.99
b (Å)	79.02	100.79	81.14
c (Å)	117.89	120.11	117.80
β (deg)	90	90	90
detector	ADSC Q315	ADSC Q315	R-AXIS IV^{++}
wavelength (Å)	0.9792	0.9792	1.5418
resolution (Å)	50.00−1.85	50.00−1.53	34.50−1.95
high-resolution shell (Å)	1.88−1.85	1.56−1.53	2.06−1.95
total no. of observations	73879	264165	101622
no. of unique reflections	19903	165103	17887
redundancy[a]	3.7 (3.7)	1.6 (1.6)	5.7 (5.7)
average I/σ[a]	6.8 (4.1)	22.5 (2.2)	8.6 (2.2)
completeness[a] (%)	99.6 (99.6)	94.1 (98.1)	100.0 (100.0)
R_{merge}[a,b] (%)	7.2 (37)	4.2 (46)	6.0 (35)

[a]Values within parentheses refer to data for the high-resolution shell. [b]$R_{merge} = \sum|I_{hkl} - \langle I\rangle|/\sum\langle I\rangle$, where $\langle I\rangle$ is the average of individual measurements of I_{hkl}.

dx.doi.org/10.1021/bi401000e | Biochemistry 2013, 52, 8048−8059

Table 3. Crystallographic Refinement Statistics

	LEA (high pH)	LEA (low pH)	LEA-Q42A
PDB entry	4DXN	4GOB	4DXQ
resolution range (Å)	41.50−1.85	33.98−1.53	30.88−1.95
no. of reflections	18802	134936	17266
R_{cryst} (%)[a]	15.1	17.5	18.5
R_{free} (%)[a]	19.5	23.1	23.8
rmsd for bond lengths (Å)	0.014	0.011	0.022
rmsd for bond angles (deg)	1.52	1.48	1.82
average B factor (Å^2)	23	24	29
total no. of atoms	1905	7502	1806
residues modeled			
chain A	4−219	2− 220	3−219
chain B		3−222	
chain C		2−220	
chain D		3−221	

[a] R_{cryst} and $R_{free} = \sum_h(||F(h)_{obs}| - |F(h)_{calc}||)/|F(h)_{obs}|$ for reflections in the working and test sets (5% of all data).

substitutions identified to correlate with photoconversion results in unstructuring of the C-terminal clasp, a feature that appears to be common to all Kaede-type proteins. To determine whether a loosening of the C-terminal clasp weakens subunit interactions, the oligomerization state of LEA in solution was determined by size-exclusion HPLC and static light scattering. The data indicate that the oligomerization state is highly concentration-dependent. When 5 mg/mL protein was injected, ∼90% of LEA eluted as a tetramer, with ∼10% higher-order aggregates. However, when 0.25 mg/mL protein was injected, only ∼20% eluted as a tetramer, whereas roughly 70% eluted as a dimer. Although these values may not represent full equilibration, the tetrameric species appears to dissociate with relative ease, in line with the crystallographically observed reduction in the number of subunit−subunit interactions.

Coplanar Preorganization Is Disrupted by the Non-photoconvertible LEA-Q42A Variant. The LEA crystallographic model demonstrates that a high degree of coplanarity exists between the GFP-like chromophore and the side chain of residue 65 (Figure 4A). Although His65(62) is not π-conjugated to the green chromophore because of light protection of the crystals, the imidazole ring is found to be arranged nearly perfectly for optimized π-overlap upon oxidation. As noted previously,[7,13,16] no major structural rearrangements have been observed upon generation of the red form. Evidently, the protein fold facilitates the extension of the chromophore skeleton by preorganizing all protein groups from which the red chromophore is formed. The $2F_O - F_C$ electron density of the chromophore and its immediate environment suggests that coplanarity may be maintained by Gln42(38), because its side chain amide is found to be π-stacked on top of the imidazole ring of His65(62) (Figure 4A). To investigate this notion, Gln42(38) was replaced with Ala, the variant was crystallized in space group $I222$ with one molecule per asymmetric unit (Table 2), and the X-ray structure of the green form of LEA-Q42(38)A was determined to 1.95 Å (Table 3). We found that this amino acid substitution disrupts the coplanarity of ring systems by favoring a different His65(62) side chain rotamer (Figure 5). Notably, a complete loss of G/R photoconversion is observed in this variant,

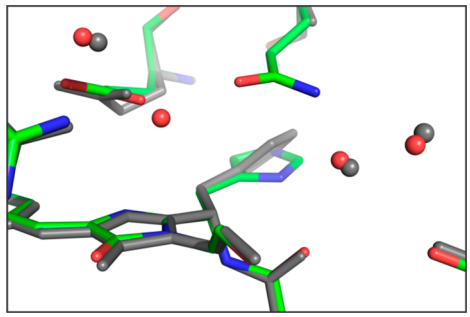


Figure 5. In the G/R photoconvertible LEA protein (carbon atoms colored green), the imidazole ring of His65(62) is in-plane with the green chromophore. In the nonphotoconvertible LEA-Q42A protein (gray), the imidazole ring has rotated to fill the vacated space, and coplanarity with the green chromophore is lost.

suggesting that structural preorganization of the two ring systems is critical for the generation of the red form.

A Hydrogen-Bonded Network Appears To Connect Glu222(211) with Phe64(61). In the LEA X-ray model, the carboxylate of Glu222(211) was found to be positioned within van der Waals distance of the His65(62) β-carbon [3.5 Å (Figure 4)], consistent with previous reports. This residue is thought to abstract a proton from the β-carbon, thus facilitating the full conjugation of the imidazole ring with the GFP-like chromophore (Scheme 1). Surprisingly, a hydrogen-bonded network that leads from the Glu222(211) carboxylate to the Phe64(61) amide oxygen could be identified (Figure 4B). During C_α−N bond scission, the amide connecting residues 64 and 65 must be ejected, forming the iminol tautomer of the product species upon protonation (Scheme 1). Hydrogen bonding from Glu222(211) to the Gln42(38) amide oxygen, from the Glu42(38) amide nitrogen to a water molecule, and from the water to Phe64(61) suggests the possibility of long-range proton transfer, such that a proton abstracted from His65(62) C_β could be delivered to the carbonyl oxygen of Phe64(61) (Scheme 1, path A). Although, in general, amide−iminol tautomerizations are energetically costly, this idea cannot be excluded on the basis of recent evidence in unrelated enzyme systems (see Discussion).

Notably, all H-bonding interactions observed in LEA that involve the Gln42(38) side chain adopt nearly ideal geometries, a feature that could facilitate long-range proton transfer via a Grotthuss (proton hopping) mechanism. To test this idea, substitutions Q42A, Q42N, and Q42M were introduced into LEA. In all three variants, photoconversion was completely abolished, indicating that Gln42(38) plays a critical role either in structural preorganization or in catalysis. As stated above, the 1.9 Å LEA-Q42A X-ray structure demonstrated that the imidazole ring of His65(62) has rotated into the vacated space previously occupied by the side chain amide of Gln42(38) (Figure 5). Therefore, the lack of photoconversion in this variant could result either from a nearly perpendicular arrangement of ring systems or from a disrupted hydrogen bond network that prevents proton transfer. Unfortunately, the crystals obtained for the LEA-Q42N and LEA-Q42M variants were not of diffraction-quality; hence, distinguishing between these two possibilities will require further experimentation.

DISCUSSION

GFP-like proteins serve as an important model system for studying fundamental processes in photobiology, such as

dx.doi.org/10.1021/bi401000e | Biochemistry 2013, 52, 8048−8059

Förster resonance energy transfer (FRET) to acceptor chromophores, excited-state proton transfer (ESPT) involving hydrogen-bonded networks, photoswitching between light and dark states by means of *cis*−*trans* isomerization, and light-induced bond rearrangements that may involve oxidation or β-elimination reactions. Additionally, GFP-like proteins have allowed us to gain substantial insight into the variety of biosynthetic mechanisms employed by nature to generate amino acid-derived cofactors. Although it has long been known that the protein scaffold provides the catalytic groups for GFP chromophore synthesis, the biochemical and photophysical features responsible for full fluorophore maturation are not always well understood. In photochemical reactions that proceed with a low quantum yield, the limited mechanistic understanding may in large part be due to the difficulty of direct observation of intermediate states.

Here, we present a novel approach to studying the structure−mechanism relationship for an energetically difficult process involving abstraction of a proton from a carbon acid concomitant with main chain cleavage. The high-resolution X-ray structure of LEA has allowed us, for the first time, to delineate the structural features critical for the irreversible G/R photoconversion process. LEA is a protein entirely constructed in the laboratory that bears the minimal set of residue substitutions essential for facile photochemical conversion. We have demonstrated that the LEA photochemistry proceeds with a maximal quantum yield of 1.5×10^{-3}, similar to the quantum yield reported for extant G/R proteins (2.4×10^{-4} for Kaede[1] and 4.7×10^{-3} for KikGR[14]). Further, we have demonstrated that the photoconversion rate depends on two apparent pK_a values, 4.5 and 7.5, rather than just one as previously suggested. Although a number of crystal structures have been published for Kaede-type proteins, to date, a structural characterization of residues known to synergistically promote color modification has not been reported. Here, we show that the substitutions responsible for efficient photoconversion consist of the replacement of large residues with smaller ones within the chromophore binding pocket, in addition to several C-terminal replacements that appear to disrupt C-tail interactions with the adjacent protomer in tetrameric assemblies. Although LEA remains tetrameric at higher concentrations, size-exclusion chromatography indicates that dimeric species dominate at low protein concentrations. Similar C-terminal disorder has previously been described for some photoswitchabe FPs.[34−36] In work on Dronpa, a loosening of the C-terminal clasp has been correlated with increased structural flexibility of the β-strands near the chromophore's phenolic end, such that dynamic motions in this part of the β-barrel promote photochromism.[37] In combination, our observations raise the tantalizing possibility that protein conformational dynamics may also facilitate light-induced chemical changes in Kaede-type proteins. In what follows, we present a hypothetical mechanism by which light-triggered conformational adjustments allow for bond rearrangements that are controlled by the electrostatics in the chromophore binding pocket.

A Light-Triggered Disruption of the Glu222−His203 Ionic Bond May Activate the Carboxylate toward Proton Abstraction. The protonation equilibria of Glu222(211), His203(193), Glu148(144), and Arg69(66) are expected to be electrostatically linked because of the alternating arrangement of charges in the chromophore binding pocket (Figure 4C). Here, we argue that the charge stabilization of the Glu222(211) carboxylate must be disrupted for it to fulfill its role as a base in the abstraction of a proton from the C_β atom of His65(62). We postulate that chromophore light absorption provides the necessary energy to transiently restructure the active site by facilitating chromophore twisting motions around its β-methylene bridge (Scheme 1, step 1). Continued irradiation with 405 nm light may lead to the buildup of a steady-state population of nonplanar chromophore conformations, resulting in a decrease in A-band absorbance. In our kinetic analysis, this effect may be attributed to the fast phase of intensity loss (Figure 3A). We speculate that chromophore motions related to *cis*−*trans* isomerization may break the salt bridge between Glu222(211) and His203(193), which is π-stacked over the chromophore's phenolic end in its equilibrium position (Figure 4). According to this model, light energy is utilized to disrupt ionic contacts, thereby transiently increasing the pK_a value of the Glu222(211) carboxylate. This process would endow the carboxylate with substantially increased proton affinity, such that abstraction of a proton from a carbon acid becomes energetically feasible.

Concerted active site remodeling in response to chromophore distortions has previously been demonstrated for the photoswitchable proteins mTFP0.7 and Dronpa.[34,35] In the dark-adapted state of these FPs, the stacking interaction between His203 and the chromophore is lost, and the guanidinium group of Arg69(66) is repositioned by a series of side chain rotations. To account for the observed structural changes, the ionic bond between His203 and Glu222 must be temporarily broken. In these proteins, Val165(157) has also been reported to undergo rearrangements upon chromophore isomerization,[34,35] whereas larger residues such as Ile165(157) have been shown to hinder isomerization.[31] Because in LEA, the V165(157)I substitution correlates with G/R photoconversion efficiency, we speculate that this residue serves to fine-tune active site motions to optimize the extent of rearrangements for photoconversion. The T72(69)A and Y121(116)N substitutions may play a similar role in providing space for structural adjustments in response to chromophore twisting, such that the chemical steps can proceed.

Protein Titratable Groups Control the Photoconversion Rate of the Chromophore. Our kinetic data indicate that the LEA photoconversion rate decreases both under acidic and under basic conditions (Figure 3C). We suggest that the apparent pK_a values of 4.5 and 7.5 extracted by curve fitting procedures may be attributed to the catalytic base Glu222(211) for the up-slope and His203(193) for the down-slope (Figure 4C). On other hand, it appears unlikely that the acid dissociation constant of 7.5 arises from titration of His65(62). In the LEA X-ray structure, this side chain is imbedded in a hydrophobic pocket with poor hydrogen bonding potential, rendering a pK_a depression below 6 significantly more likely than an elevation by ~1.5 pK_a units above the value established for an aqueous environment. Such a substantial increase in the apparent pK_a value would require electrostatic interactions that stabilize the imidazolium cation; however, such interactions are not observed in the structure.

The two glutamate residues, Glu222(211) and Glu148(144), must be anionic, and His203(193) and Arg69(66) must be cationic to maintain the electrostatic network prior to irradiation. The LEA chromophore pK_a value falls between the midpoints for the up- and down-slopes, suggesting that its protonation equilibrium is strongly modulated by the surrounding electrostatic field. Therefore, we predict that the

bell-shaped curve (Figure 3C) will be shifted to a higher pH range in variants with elevated chromophore pK_a values.

Is Photoconversion Linked to a Proton Relay from Glu222 to Phe64? Overall, the net reaction requires removal of a proton from the β-carbon of His65(62) and addition of a proton to the carbonyl oxygen of Phe64(61) to generate the product iminol upon bond scission. Such a scenario could be accommodated by long-range proton transfer via a hydrogen bond network leading from the donor to the acceptor (Scheme 1, path A) or by proton shuttling via an appropriate proton carrier (Scheme 1, path B). On the basis of the LEA structure presented here, one could envision that abstraction of a proton from the β-carbon of His65(62) (Scheme 1, step 2A) may be followed by proton hopping along a crystallographically observed H-bonded network that leads to the carbonyl oxygen of Phe64(61). According to this scenario, a rapid proton relay to the acceptor oxygen would compete with proton back transfer, facilitating bond scission via β-elimination by stabilizing the leaving group (step 3A). Proton transfer would result in tautomerization of the Gln42(38) side chain amide to its iminol form, with subsequent re-equilibration to the low-energy amide. In neutral aqueous solutions, the amide form is known to be the dominant tautomer ($K_{eq} = 10^{-8}$);[38] therefore, the equilibrium must be perturbed by several orders of magnitude to account for the LEA photoconversion quantum yield of 1.5×10^{-3}. This could be achieved by subtle changes in hydrogen bonding geometries that stabilize the iminol form in the protein's interior. In LEA, the amide oxygen appears to be without a H-bonding partner but is well-poised to accept a proton from Glu222(211) (Figure 4). In support of this notion, tautomerization of glutamine to its iminol form is thought to be a key step in the light activation of BLUF-domain proteins,[39−41] where the relay of a proton to the flavin group has been proposed to involve the side chain of a glutamine residue.[41−44]

Although in Kaede-type proteins, long-range proton transfer remains highly speculative, optimized H-bonding geometries as observed in LEA could play a role in this type of mechanism. Although not previously reported in the literature, a similar hydrogen-bonded network providing a proton hopping pathway from Glu222 to Phe64 is also observed in other G/R photoconvertible proteins such as Kaede (PDB entry 2GW3), EosFP (PDB entry 1ZUX), Dendra2 (PDB entry 2VZX), and KikGR (PDB entry 2DDD), underscoring its potential significance in the photochemical reaction.

Is Photoconversion Linked to Stepwise Proton Shuttling Mediated by His65 Side Chain Rotations? As an alternative to path A, we suggest that His65 may serve as a direct proton shuttle between Glu222 and Phe64 (Scheme 1, path B). Previously proposed mechanistic models have invoked transfer of a proton from the imidazolium cation of His65(62) to the carbonyl oxygen of Phe64(61).[13,16,30] Although the origin of His65(62) protonation has remained obscure, this idea is supported by the observation that aromatic residues other than histidine cannot sustain G/R photochemistry.[2,12] In LEA, direct transfer of a proton from His65(62) to Phe64(61) appears to be feasible, because Nδ1 of His65(62) is located within 3.3 Å of the carbonyl oxygen of Phe64(61), although with an unfavorable angular orientation. The replacement of the large Tyr121 with a much smaller residue [Y121(116)N] near the imidazole group has been shown to be critical for photoconversion. Residue 121 may exert its function primarily by providing space for alternate His65(62) side chain conformations populated upon photoexcitation of the chromophore.

The proton shuttling process could be facilitated by side chain rotations in conjunction with protein breathing motions similar to those described in Dronpa.[37] As noted above, excited-state chromophore deformations may activate the carboxylate of Glu222(211) toward abstraction of a proton from C_β of His65(62) (Scheme 1, step 1). As in path A, neutralization of Glu222(211) would generate a carbanion at C_β of His65(62) (step 2AB). However, along path B, this event would be followed by His65(62) C_α−C_β bond rotation (step 3B), facilitating retrieval of the proton by the imidazole Nδ1 atom from Glu222(211) (step 4B). A subsequent C_α−C_β bond rotation would reorient the imidazolium cation (step 5B) to deliver a proton to the Phe64(61) carbonyl oxygen. As the leaving group is ejected as an iminol, the C−N bond is cleaved (step 6B). The final step would consist of tautomerization to the corresponding amide group by means of solvent molecules (step 7AB). Experimental support in favor of path B is provided by a crystallographic report on the G/R photoconvertible protein KikGR, in which His65 was demonstrated to undergo a C_α−C_β bond rotation during the photochemical reaction.[15] With a separation of only 3.4 Å, the new side chain rotamer was shown to place Nδ1 of His65 in the proximity of the Glu222 carboxylate, rendering proton transfer according to step 3B (Scheme 1) geometrically feasible.

■ CONCLUSIONS

Using a least evolved ancestor protein, we have delineated the pH dependence of the formation of a red-emitting three-ring chromophore. We assign a functional role to the internal charge network in the transient activation of the base toward proton abstraction. In addition, we propose two alternate mechanisms describing the photoconversion chemistry, both triggered by excited-state chromophore twisting motions promoted by a decrease in the level of molecular packing in the binding pocket. According to the first model, a proton is relayed along a conserved H-bonded network involving Gln42, and according to the second model, histidine rotamer sampling provides a shuttle mechanism for the delivery of a proton to the acceptor group. In either case, light-induced disruption of an internal salt bridge may provide the driving force for the ensuing chemical steps.

■ ASSOCIATED CONTENT

Accession Codes

Atomic coordinates and structure factors for the LEA crystal structures have been deposited in the Protein Data Bank as entries 4DXN for the high-pH form of LEA, 4GOB for the low-pH form of LEA, and 4DXQ for LEA-Q38A.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: rwachter@asu.edu. Phone: (480) 965-8188. Fax: (480) 965-2747.

Funding

This work was supported by National Science Foundation Grant MCB-0615938 to R.M.W.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

Results presented here were in part derived from work performed at the Argonne National Laboratory Structural Biology Center at the Advanced Photon Source. Argonne National Laboratory is operated by the University of Chicago Argonne, LLC, for the U.S. Department of Energy, Office of Biological and Environmental Research, under Contract DE-AC02-06CH11357.

■ ABBREVIATIONS

FP, fluorescent protein; GFP, green fluorescent protein; LEA, least evolved ancestor; TLS, translation−libration−screw rotation; G/R, green-to-red; rmsd, root-mean-square deviation.

■ REFERENCES

(1) Ando, R., Hama, H., Yamamoto-Hino, M., Mizuno, H., and Miyawaki, A. (2002) An optical marker based on the UV-induced green-to-red photoconversion of a fluorescent protein. *Proc. Natl. Acad. Sci. U.S.A. 99*, 12651−12656.

(2) Wiedenmann, J., Ivanchenko, S., Oswald, F., Schmitt, F., Roecker, C., Salih, A., Spindler, K.-D., and Nienhaus, G. U. (2004) EosFP, a fluorescent marker protein with UV-inducible green-to-red fluorescence conversion. *Proc. Natl. Acad. Sci. U.S.A. 101*, 15905−15910.

(3) Labas, Y. A., Gurskaya, N. G., Yanushevich, Y. G., Fradkov, A. F., Lukyanov, K. A., Lukyanov, S. A., and Matz, M. V. (2002) Diversity and evolution of the green fluorescent protein family. *Proc. Natl. Acad. Sci. U.S.A. 99*, 4256−4261.

(4) Pakhomov, A. A., Martynova, N. Y., Gurskaya, N. G., Balashova, T. A., and Martynov, V. I. (2004) Photoconversion of the chromophore of a fluorescent protein from Dendronephthya sp. *Biochemistry (Moscow) 69*, 901−908.

(5) Matz, M. V., Labas, Y. A., and Ugalde, J. (2006) Evolution of function and color in GFP-like proteins. *Methods Biochem. Anal. 47*, 139−161.

(6) Lippincott-Schwartz, J., and Patterson, G. H. (2009) Photoactivatable fluorescent proteins for diffraction-limited and super-resolution imaging. *Trends Cell Biol. 19*, 555−565.

(7) Wachter, R. M., Watkins, J. L., and Kim, H. (2010) Mechanistic diversity of red fluorescence acquisition by GFP-like proteins. *Biochemistry 49*, 7417−7427.

(8) Wiedenmann, J., Gayda, S., Adam, V., Oswald, F., Nienhaus, K., Bourgeois, D., and Nienhaus, G. U. (2011) From EosFP to mIrisFP: Structure-based development of advanced photoactivatable marker proteins of the GFP-family. *J. Biophotonics 4*, 377−390.

(9) Alieva, N. O., Konzen, K. A., Field, S. F., Meleshkevitch, E. A., Hunt, M. E., Beltran-Ramirez, V., Miller, D. J., Wiedenmann, J., Salih, A., and Matz, M. V. (2008) Diversity and evolution of coral fluorescent proteins. *PLoS One 3*, e2680.

(10) Ugalde, J. A., Chang, B. S. W., and Matz, M. V. (2004) Evolution of coral pigments recreated. *Science 305*, 1433−1433.

(11) Field, S. F., and Matz, M. V. (2010) Retracing evolution of red fluorescence in GFP-like proteins from Faviina corals. *Mol. Biol. Evol. 27*, 225−233.

(12) Mizuno, H., Mal, T. K., Tong, K. I., Ando, R., Furuta, T., Ikura, M., and Miyawaki, A. (2003) Photo-induced peptide cleavage in the green-to-red conversion of a fluorescent protein. *Mol. Cell 12*, 1051−1058.

(13) Nienhaus, K., Nienhaus, G. U., Wiedenmann, J., and Nar, H. (2005) Structural basis for photo-induced protein cleavage and green-to-red conversion of fluorescent protein EosFP. *Proc. Natl. Acad. Sci. U.S.A. 102*, 9156−9159.

(14) Tsutsui, H., Karasawa, S., Shimizu, H., Nukina, N., and Miyawaki, A. (2005) Semi-rational engineering of a coral fluorescent protein into an efficient highlighter. *EMBO Rep. 6*, 233−238.

(15) Tsutsui, H., Shimizu, H., Mizuno, H., Nukina, N., Furuta, T., and Miyawaki, A. (2009) The E1 Mechanism in Photo-Induced β-Elimination Reactions for Green-to-Red Conversion of Fluorescent Proteins. *Chem. Biol. 16*, 1140−1147.

(16) Hayashi, I., Mizuno, H., Tong, K. I., Furuta, T., Tanaka, F., Yoshimura, M., Miyawaki, A., and Ikura, M. (2007) Crystallographic evidence for water-assisted photo-induced peptide cleavage in the stony coral fluorescent protein Kaede. *J. Mol. Biol. 2007*, 918−926.

(17) Lelimousin, M., Adam, V., Nienhaus, G. U., Bourgeois, D., and Field, M. J. (2009) Photoconversion of the fluorescent protein EosFP: A hybrid potential simulation study reveals intersystem crossings. *J. Am. Chem. Soc. 131*, 16814−16823.

(18) Li, X., Chung, L. W., Mizuno, H., Miyawaki, A., and Morokuma, K. (2010) Competitive mechanistic pathways for green-to-red photoconversion in the fluorescent protein Kaede: A computational study. *J. Phys. Chem. B 114*, 16666−16675.

(19) Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol. 276*, 307−326.

(20) McCoy, A., Grosse-Kunstleve, R., Adams, P., Winn, M., Storoni, L., and Read, R. (2007) Phaser crystallographic software. *J. Appl. Crystallogr. 40*, 658−674.

(21) Cowtan, K. (2006) The Buccaneer software for automated model building. *Acta Crystallogr. D62*, 1002−1011.

(22) Emsley, P., and Cowtan, K. (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D60*, 2126−2132.

(23) Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelyhood method. *Acta Crystallogr. D53*, 240−255.

(24) Nienhaus, K., Renzi, F., Vallone, B., Wiedenmann, J., and Nienhaus, G. U. (2006) Exploring chromophore-protein interactions in fluorescent protein cmFP512 from *Cerianthus membranaceus*: X-ray structure analysis and optical spectroscopy. *Biochemistry 45*, 12942−12953.

(25) Elsliger, M.-A., Wachter, R. M., Hanson, G. T., Kallio, K., and Remington, S. J. (1999) Structural and spectral response of green fluorescent protein variants to changes in pH. *Biochemistry 38*, 5296−5301.

(26) Painter, J., and Merritt, E. A. (2006) Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr. D62*, 439−450.

(27) Leslie, A. G. (1999) Integration of macromolecular diffraction data. *Acta Crystallogr. D55*, 1696−1702.

(28) Collaborative Computational Project No. 4 (1994) The CCP4 Suite: Programs for Protein Crystallography. *Acta Crystallogr. D50*, 760−763.

(29) Segel, I. H. (1975) in *Enzyme kinetics: Behavior and analysis of rapid equilibrium and steady state enzyme systems* (Segel, I. H., Ed.) Wiley, New York.

(30) Adam, V., Nienhaus, K., Bourgeois, D., and Nienhaus, G. U. (2009) Structural basis of enhanced photoconversion yield in green fluorescent protein-like protein Dendra2. *Biochemistry 48*, 4905−4915.

(31) Adam, V., Moeyaert, B., David, C. C., Mizuno, H., Lelimousin, M., Dedecker, P., Ando, R., Miyawaki, A., Michiels, J., Engelborghs, Y., and Hofkens, J. (2011) Rational design of photoconvertible and biphotochromic fluorescent proteins for advanced microscopy applications. *Chem. Biol. 18*, 1241−1251.

(32) Seemann, K. M., Kiefersauer, R., Jacob, U., and Kuhn, B. (2012) Optical pH detection within a protein crystal. *J. Phys. Chem. B 116*, 9873−9881.

(33) Yarbrough, D., Wachter, R. M., Kallio, K., Matz, M. V., and Remington, S. J. (2001) Refined crystal structure of DsRed, a red fluorescent protein from coral, at 2.0-Å resolution. *Proc. Natl. Acad. Sci. U.S.A. 98*, 462−467.

(34) Henderson, J. N., Ai, H.-w., Campbell, R. E., and Remington, S. J. (2007) Structural basis for reversible photobleaching of a green fluorescent protein homologue. *Proc. Natl. Acad. Sci. U.S.A. 104*, 6672−6677.

(35) Andresen, M., Stiel, A. C., Trowitzsch, S., Weber, G., Eggeling, C., Wahl, M. C., Hell, S. W., and Jakobs, S. (2007) Structural basis for

reversible photoswitching in Dronpa. *Proc. Natl. Acad. Sci. U.S.A. 104*, 13005−13009.
(36) Stiel, A. C., Trowitzsch, S., Weber, G., Andresen, M., Eggeling, C., Hell, S. W., Jakobs, S., and Wahl, M. C. (2007) 1.8 Å bright-state structure of the reversibly switchable fluorescent protein Dronpa guides the generation of fast switching variants. *Biochem. J. 402*, 35−42.
(37) Mizuno, H., Mal, T. K., Walchli, M., Kikuchi, A., Fukano, T., Ando, R., Jeyakanthan, J., Taka, J., Shiro, Y., Ikura, M., and Miyawaki, A. (2008) Light-dependent regulation of structural flexibility in a photochromic fluorescent protein. *Proc. Natl. Acad. Sci. U.S.A. 105*, 9227−9232.
(38) Sigel, H., and Martin, R. B. (1982) Coordinating properties of the amide bond. Stability and structure of metal ion complexes of peptides and related ligands. *Chem. Rev. 82*, 385−426.
(39) Masuda, S., and Bauer, C. E. (2002) AppA is a blue light photoreceptor that antirepresses photosynthesis gene expression in *Rhodobacter sphaeroides*. *Cell 110*, 613−623.
(40) Laan, W., Gauden, M., Yeremenko, S., van Grondelle, R., Kennis, J. T., and Hellingwerf, K. J. (2006) On the mechanism of activation of the BLUF domain of AppA. *Biochemistry 45*, 51−60.
(41) Stelling, A. L., Ronayne, K. L., Nappa, J., Tonge, P. J., and Meech, S. R. (2007) Ultrafast structural dynamics in BLUF domains: Transient infrared spectroscopy of AppA and its mutants. *J. Am. Chem. Soc. 129*, 15556−15564.
(42) Domratcheva, T., Grigorenko, B. L., Schlichting, I., and Nemukhin, A. V. (2008) Molecular models predict light-induced glutamine tautomerization in BLUF photoreceptors. *Biophys. J. 94*, 3872−3879.
(43) Sadeghian, K., Bocola, M., and Schutz, M. (2008) A conclusive mechanism of the photoinduced reaction cascade in blue light using flavin photoreceptors. A QM/MM study on the fast photocycle of blue light using flavin photoreceptors in their light-adapted/active form. *J. Am. Chem. Soc. 130*, 12501−12513.
(44) Sadeghian, K., Bocola, M., and Schutz, M. (2010) A QM/MM study on the fast photocycle of blue light using flavin photoreceptors in their light-adapted/active form. *Phys. Chem. Chem. Phys. 12*, 8840−8846.