The Case of the Missing Ring: Radical Cleavage of a Carbon−Carbon Bond and Implications for GFP Chromophore Biosynthesis

ARTICLE *in* JOURNAL OF THE AMERICAN CHEMICAL SOCIETY · APRIL 2007

Impact Factor: 12.11 · DOI: 10.1021/ja063983u · Source: PubMed

4 AUTHORS, INCLUDING:

David P Barondeau
Texas A&M University
29 PUBLICATIONS **1,132** CITATIONS
SEE PROFILE

John Tainer
University of Texas MD Anderson Cancer Center
456 PUBLICATIONS **30,090** CITATIONS
SEE PROFILE

Understanding GFP Chromophore Biosynthesis: Controlling Backbone Cyclization and Modifying Post-translational Chemistry[†,‡]

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

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

Received September 24, 2004; Revised Manuscript Received November 17, 2004

ABSTRACT: The *Aequorea victoria* green fluorescent protein (GFP) undergoes a remarkable post-translational modification to create a chromophore out of its component amino acids S65, Y66, and G67. Here, we describe mutational experiments in GFP designed to convert this chromophore into a 4-methylidene-imidazole-5-one (MIO) moiety similar to the post-translational active-site electrophile of histidine ammonia lyase (HAL). Crystallographic structures of GFP variant S65A Y66S (GFPhal) and of four additional related site-directed mutants reveal an aromatic MIO moiety and mechanistic details of GFP chromophore formation and MIO biosynthesis. Specifically, the GFP scaffold promotes backbone cyclization by (1) favoring nucleophilic attack by close proximity alignment of the G67 amide lone pair with the π^* orbital of the residue 65 carbonyl and (2) removing enthalpic barriers by eliminating inhibitory main-chain hydrogen bonds in the precursor state. GFP R96 appears to induce structural rearrangements important in aligning the molecular orbitals for ring cyclization, favor G67 nitrogen deprotonation through electrostatic interactions with the Y66 carbonyl, and stabilize the reduced enolate intermediate. Our structures and analysis also highlight negative design features of the wild-type GFP architecture, which favor chromophore formation by destabilizing alternative conformations of the chromophore tripeptide. By providing a molecular basis for understanding and controlling the driving force and protein chemistry of chromophore creation, this research has implications for expansion of the genetic code through engineering of modified amino acids.

To achieve their biological function, proteins are driven by evolutionary selection to augment the chemical reactivity and properties of natural amino acids through post-translation modifications (*1*). Although many of these modifications require additional enzymes, some are self-catalyzed, including the tripeptide backbone cyclization reactions that create the fluorophore of green fluorescent protein (GFP)[1] (*2*, *3*) and the electrophilic catalysts (*4*−*6*) of the enzymes histidine ammonia lyase (HAL) (*7*), phenylalanine ammonia lyase (PAL) (*8*), and tyrosine aminomutase (TAM) (*9*). Similarities in both the resulting five-membered imidazolone moieties (Figure 1) and their biosynthetic reaction mechanisms, despite dramatically different protein environments and architectures (Figure 2), suggest opportunities to engineer, modify, and transplant post-translational chemistry between protein scaffolds. Determining the driving forces and molecular mechanisms of these natural systems will enhance our ability to design and control such self-catalyzed amino acid transformations, thus effectively expanding the genetic code for novel applications in biology, biotechnology, and

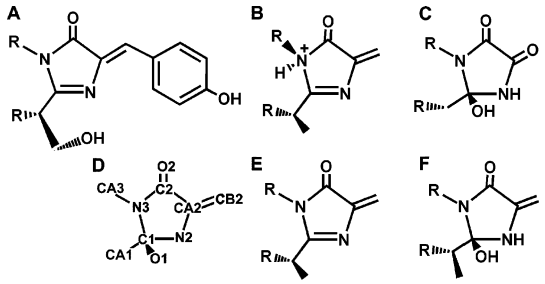

FIGURE 1: Molecular structures of imidazolone products after backbone cyclization and other post-translational modifications. (A) aromatic GFP chromophore, (B) nonaromatic HAL MIO, (C) nonaromatic GFP Gly-Gly-Gly, (D) atom labels for GFPhal H148G variant, (E) modified aromatic chromophore for GFPhal variant, and (F) nonaromatic moiety for the GFPhal H148G variant. In each case, the N-terminal polypeptide chain enters the imidazolone ring at the bottom left and exits C-terminally from the top left.

[†] This work was supported by the La Jolla Interfaces in Sciences (to D.P.B.), NIH GM19290 Postdoctoral Fellowships (to D.P.B.), and NIH Grant RO1 GM37684 (to E.D.G.). Portions of this research were carried out at the Stanford Synchrotron Radiation Laboratory, a national user facility operated by Stanford University on behalf of the U.S. Department of Energy, Office of Basic Energy Sciences. The SSRL Structural Molecular Biology Program is supported by the Department of Energy, Office of Biological and Environmental Research, the National Institutes of Health, National Center for Research Resources, Biomedical Technology Program, and the National Institute of General Medical Sciences.

[‡] Coordinates and structure factors for all reported data sets have been deposited in the Protein Data Bank (PDB entries 1YJF, 1YJ2, 1YHI, 1YHH, and 1YHG).

* To whom correspondence should be addressed. Telephone: 858-784-2878. Fax: 858-784-2289. E-mail: edg@scripps.edu.

[1] Abbreviations: GFP, green fluorescent protein; MIO, 4-methylidene-imidazole-5-one; HAL, histidine ammonia lyase; GFPuv, S65T F99S M153T V163A GFP variant; GFPsol, F64L GFPuv variant; GFPhal, S65A Y66S GFPuv variant; PAL, phenylalanine ammonia lyase; TAM, tyrosine aminomutase; RFP, red fluorescent protein; Gly-Gly-Gly, S65G Y66G GFP variant; PDB, Protein Data Bank; rmsd, root-mean-square deviation; MW, molecular weight.

10.1021/bi0479205 CCC: $30.25
Published on Web 01/22/2005

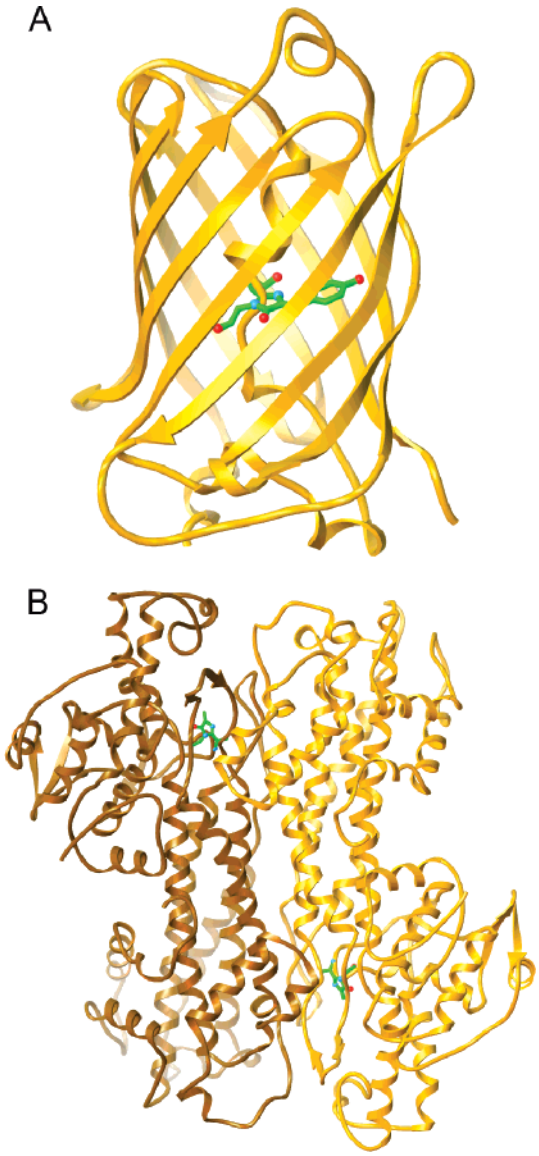

FIGURE 2: Overall protein fold and location of the post-translational modifications for (A) GFP and (B) HAL. Post-translational cyclization products for both GFP and HAL are shown in green. The two subunits of the HAL homodimer are displayed in brown and yellow.

medicine. For biology and biotechnology, the spontaneous biosynthesis (*10*, *11*) and tunable fluorescence properties of the chromophores of GFP (*2*, *3*), its homologues (*12*, *13*), and mutants (*14*, *15*), have revolutionized *in vivo* molecular tagging and cell labeling. For medicine, deficiencies in HAL activity, the nonoxidative elimination of the α-amino group of histidine, lead to the rare hereditary metabolic disorder histidinemia that is often characterized by mental retardation and speech defects (*16*); PAL may be used to degrade phenylalanine and combat the metabolic disease phenylketonuria (*17*); and TAM is a key enzyme in enediyne antitumor antibiotic C-1027 biosynthesis (*9*).

Both GFP fluorophore and HAL 4-methylidene-imidazole-5-one (MIO) post-translational modifications entail three major synthetic steps: backbone cyclization via covalent bond formation between glycine nitrogen (G67 in GFP and G144 in HAL) and carbonyl carbon atoms (S65 in GFP and A142 in HAL), dehydration of the oxygen of the same carbonyl, and either Y66 oxidation (GFP) (*14*, *18*, *19*) or S143 dehydration (HAL) (*7*, *20*) reactions to generate Cα−Cβ double bonds and mature ring systems. The order of the three reactions in GFP is under debate, and proposals include cyclization−dehydration−oxidation (*18*), cyclization−oxidation−dehydration (*21*), and oxidation−cyclization−dehydration (*22*) mechanisms (Figure 3). Some GFP homologues [e.g., red fluorescent protein (RFP)] extend the electronic conjugation of the GFP-like chromophore with a fourth major step, a second oxidation reaction that generates a double bond between backbone nitrogen and Cα atoms (*23*−*25*). Interestingly, a major difference between the structures of the imidazolone rings generated through the post-translational modification of GFP/RFP and HAL/PAL is the hybridization of the N3 atom and the resulting aromaticity for the fluorophore but not the MIO moiety (Figure 1).

The cyclization for fluorophore and MIO biosynthesis is an electronically unfavorable reaction (*20*), and three different proposals have emerged to explain how the proteins overcome this apparent thermodynamic barrier. First, a mechanical compression hypothesis was suggested, in which peptide cyclization is driven by a relaxing strain imposed in the precyclization state by the protein architecture of GFP (*26*) or HAL (*20*). This hypothesis is consistent with nonviable HAL mutants (*20*) that indicate that steric interactions are important for peptide cyclization. Second, interpretation of density functional theory calculations led to proposals that side-chain oxidation (GFP) (*22*) or dehydration (HAL) (*27*) occurs first and drives backbone cyclization. Third, a recent conjugation-trapping mechanism for GFP was proposed, supported by precyclization and postcyclization structures for the R96A and S65G Y66G (Gly-Gly-Gly) variants, in which the initial thermodynamically unfavorable cyclization event is coupled to a subsequent trapping step under Le Chatliers' principle (*28*). Under aerobic but not anaerobic conditions, the post-translational product of this Gly-Gly-Gly variant is a cyclized but nonaromatic moiety that contains oxygen atoms attached to the S65G carbonyl carbon (carbonyl oxygen normally lost during dehydration) and the Y66 Cα (probably incorporated through an oxidation reaction) atoms (Figure 1C) (*28*). Thus, in contrast to HAL MIO biosynthesis, GFP fluorophore maturation and stabilization of the backbone cyclization product for chromophore variants appear to require an oxidation reaction.

The GFP and HAL post-translational modifications occur in dramatically different protein scaffolds, and their local protein environments exhibit few evident common features to assist in backbone cyclization. Extensive GFP mutagenesis of the chromophore residues and surrounding amino acids generally produces chromophores with altered spectral properties rather than incomplete maturation (*14*, *15*, *29*), implicating a major role for the protein architecture in chromophore biosynthesis and establishing GFP as a robust experimental system. Mutations at R96 and G67 are two exceptions. The mutation of R96 greatly retards GFP chromophore maturation through loss of steric and electrostatic interactions (*28*), and the mutation of G67 results in a nonfluorescent, presumably noncyclized, product (*10*). In contrast, HAL MIO biosynthesis lacks an R96 analogue (*7*), is disrupted by shortening adjacent side chains (*20*, *30*), and can occur with an Ala substitution for the G67 equivalent (mutant exhibits 37% wild-type activity) (*30*). Thus, although the initial cyclization reaction and cyclized ring products of

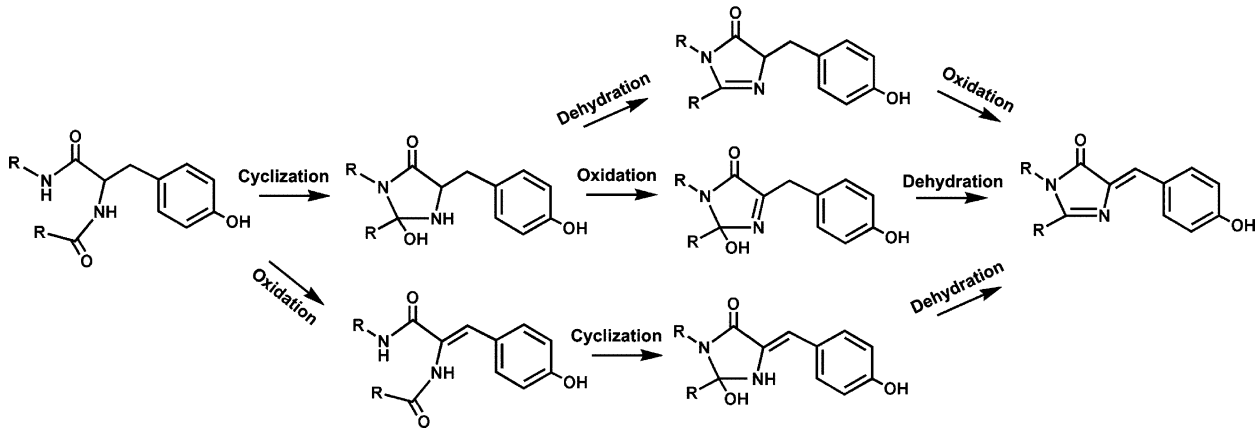

FIGURE 3: Proposed cyclization−dehydration−oxidation (top), cyclization−oxidation−dehydration (middle), and oxidation−cyclization−dehydration (bottom) reaction schemes are displayed for GFP chromophore maturation.

GFP and HAL appear similar, they occur in distinct protein scaffolds with dissimilar secondary structure, solvent accessibility, and surrounding residues that may assist in backbone cyclization.

To better understand the driving force for post-translational modification in GFP and HAL and to determine the scaffold-specific features that contribute to backbone cyclization, we transplanted the HAL electrophile sequence and then engineered additional site-directed mutations into the GFP scaffold. Our high-resolution structural data reveal architectural and mechanistic requirements for GFP peptide backbone cyclization, support and build upon previous backbone cyclization results and hypotheses (*28*), and provide an alternative explanation for modified chromophore variants, previously used to support the cyclization−oxidation−dehydration mechanism (*21*). Moreover, we determine that the S65A Y66S (GFPhal) variant forms an aromatic MIO species in the GFP scaffold, compare the GFPhal and HAL post-translational modifications, and discuss functional implications.

MATERIALS AND METHODS

Experimental Preparation, Crystallization, and Structure Refinement. Using the QuikChange method (Stratagene), we created GFPhal by replacing the GFP Ser-Tyr-Gly fluorophore residues with those of the HAL Ala-Ser-Gly electrophile. GFPhal, GFPhal H148G, and GFPhal R96A mutants were engineered from the high solubility GFPuv (S65T F99S M153T V163A) plasmid (Clontech), and the GFPhal G67A and GFPhal S65G V68G mutants from GFPsol (GFPuv with an additional F64L mutation to further enhance solubility). We purified and crystallized the proteins as previously described (*2*, *28*, *31*). Interestingly, both the GFPhal and GFPhal H148G variants exhibit heat/denaturation-induced backbone cleavage, reminiscent of RFP (*23*). X-ray diffraction data sets were collected at the Stanford Synchrotron Radiation Laboratory on beamlines 11-1, GFPhal H148G (λ = 0.85 Å), and 9-1, GFPhal (λ = 0.97 Å), GFPhal R96A (λ = 0.97 Å), GFPhal S65G V68G (λ = 0.97975 Å), and GFPhal G67A (λ = 0.97975 Å). Data sets were indexed and reduced with the *hkl* package (*32*), and phases were determined by molecular replacement with AmoRe (*33*). The search model was a refined 1.0 Å GFPsol structure, determined by molecular replacement from a previous GFP structure (*2*). Difference electron density and omit maps were manually fit with the XtalView package (*34*) and refined in either CNS (*35*) or Shelx-97 (*36*) using all of the diffraction data, except for 5% used for R_{free} calculations (*37*). Standard uncertainties were determined by inverting the full least-squares covariance matrix in Shelx-97 (*36*). All structures were superimposed with Sequoia (*38*).

RESULTS

Transplanting the HAL Electrophile Sequence into GFP. We created GFPhal (S65A Y66S) to test if the GFP protein scaffold supports MIO electrophile synthesis and to further elucidate the architectural and protein chemistry requirements for the backbone cyclization reactions. The GFPhal absorbance maximum at 385 nm (Figure 4A) indicates that post-translational modification has occurred but has created a chromophore differing from the fluorophore of both wild-type and GFPsol parents (absorbance maxima of 395 and 489 nm, respectively). The crystallographic structure (Table 1) of the essentially nonfluorescent GFPhal reveals no significant differences compared to GFPsol [overall Cα root-mean-square deviation (rmsd) of 0.20 Å], except for changes at the mutated chromophore residues and the rotation of the H148 side chain to occupy the cavity created by the Y66S truncation. The electron density for the modified chromophore reveals a five-membered ring system but no side-chain hydroxyl group or aldehyde, which would be expected from a wild-type Cα−Cβ oxidation at residue 66 (Figure 4B). The product (Figure 1E) is consistent with deprotonation of the Y66S Cα and double-bond formation by elimination of the hydroxyl group. Interestingly, the short (2.27 ± 0.07 Å) distance from the exocyclic, presumably Y66S CB2 atom to a water molecule (Wat1) coplanar with the imidazolone ring suggests an unusual CH−O hydrogen bond. An alternative explanation for these data, which cannot be ruled out by the electron density maps, is Y66S side-chain elimination followed by an oxygen incorporation reaction (see the Discussion). A second water molecule (Wat2) forms a hydrogen bond to the V68 backbone nitrogen atom (Figure 4D) and is in van der Waals contact (2.77 ± 0.08 Å) with the S65A carbonyl carbon atom. Interestingly, omit electron density for Wat2 extends to the chromophore, indicating a small population (estimate of <10%) of bound hydroxyl (see below). These two water molecules, Wat1 and Wat2, may result from the two dehydrations during MIO formation in GFP.

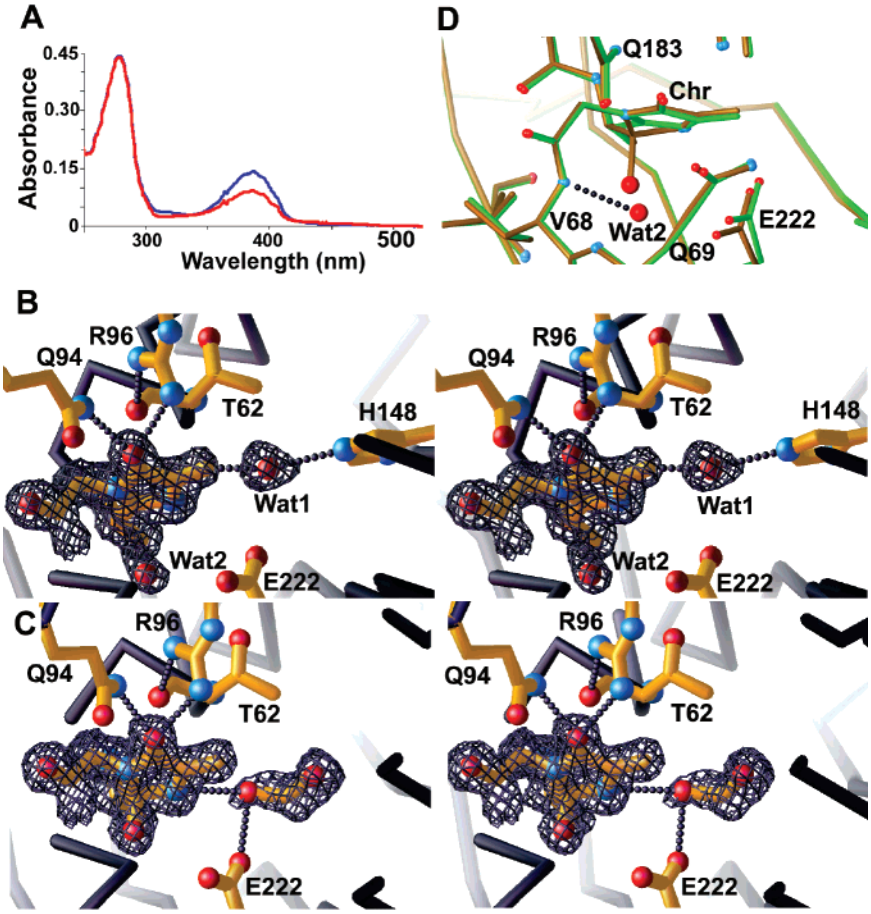

FIGURE 4: Crystallographic structures and UV−vis spectral properties displayed for the GFPhal and GFPhal H148G variants. (A) UV−vis spectral properties for GFPhal (blue) and GFPhal H148G (red) variants. (B) Stereopair of the GFPhal variant with simulated annealing omit $|F_o - F_c|$ electron density for the chromophore contoured at 3σ (black). (C) Stereopair of the GFPhal H148G GFP variant with simulated annealing omit $|F_o - F_c|$ electron density for the chromophore and ethylene glycol (cryosolvent) molecule contoured at 3σ (black). (D) Overlay of GFPhal (green) and GFPhal H148G (yellow) structures.

Table 1: Data Collection and Refinement Statistics

	GFPhal	GFPhal H148G	GFPhal R96A	GFPhal G67A	GFPhal S65G V68G
resolution (Å)	50.0−1.35	30.0−1.50	50.0−1.90	20.0−1.50	50.0−2.50
last shell (Å)[a]	1.40−1.35	1.55−1.50	1.97−1.90	1.55−1.50	2.59−2.50
space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_1$
molecules/asymmetric unit	1	1	1	1	2
unit cell	51.0, 62.6, 71.5	51.00, 62.43, 71.45	54.5, 64.3, 70.5	51.3, 62.7, 70.2	45.3, 71.1, 60.9, $\beta = 94.8$
observations	119 781	99 625	65 585	145 671	60 262
unique	49 525	36 704	19 956	35 794	13 380
R_{sym} (%)[b]	4.7 (33.8)	6.1 (37.2)	5.9 (31.4)	6.9 (35.9)	11.9 (31.2)
completeness (%)	97.1 (96.0)	98.0 (91.9)	98.6 (99.5)	96.9 (97.5)	99.4 (99.7)
$I/\sigma I$	21.1 (2.7)	16.1 (2.8)	21.8 (4.30	21.1 (3.9)	17.0 (6.5)
refinement parameter	19 851	8641	7636	8176	14 692
R_{work}/R_{free} (%)[c]	14.5/20.5	17.1/22.6	22.5/25.7	21.7/22.8	21.4/28.0
PDB code	1YJF	1YJ2	1YHI	1YHH	1YHG

[a] Values in parentheses are the statistics for the highest resolution shell of data. [b] $R_{sym} = \sum|I_{hkl} - \langle I\rangle|/\sum\langle I\rangle$, where $\langle I\rangle$ is the average individual measurement of I_{hkl}. [c] $R_{work} = (\sum|F_{obs} - F_{calc}|)/\sum|F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure factors, respectively.

The 1.35 Å resolution of the GFPhal structure allows assignment of the tautomer and primary resonance form of the MIO moiety by accurate determination of bond lengths between the non-hydrogen atoms. The MIO moiety is planar with the exocyclic CA1, O2, CA3, and CB2 atoms in the plane of the five-membered ring, indicating that the C1, C2, CA2, and N3 atoms are sp^2-hybridized (see labels in Figure 1D). The C1−CA1 (1.53 ± 0.06 Å), C1−N3 (1.36 ± 0.05 Å), and C1−N2 (1.28 ± 0.05 Å) bond lengths are consistent with a C1−N2 double bond, while the CA2−N2 (1.37 ± 0.05 Å), CA2−C2 (1.44 ± 0.06 Å), and CA2−CB2 (1.32 ± 0.06 Å) bond lengths imply a CA2−CB2 double bond. Further, the C2−O2 bond length (1.25 ± 0.06 Å) is consistent with a keto rather than enol tautomer, and the N3−C2 (1.38 ± 0.06 Å) and N3−CA3 (1.40 ± 0.04 Å) bond lengths suggest a neutral sp^2-hybridized N3 atom (Figure 1E). In contrast, the HAL MIO contains an sp^3-hybridized N3 atom (Figure 1B). Thus, the primary difference between the GFPhal MIO and the HAL MIO appears to be hybridization at the N3 atom.

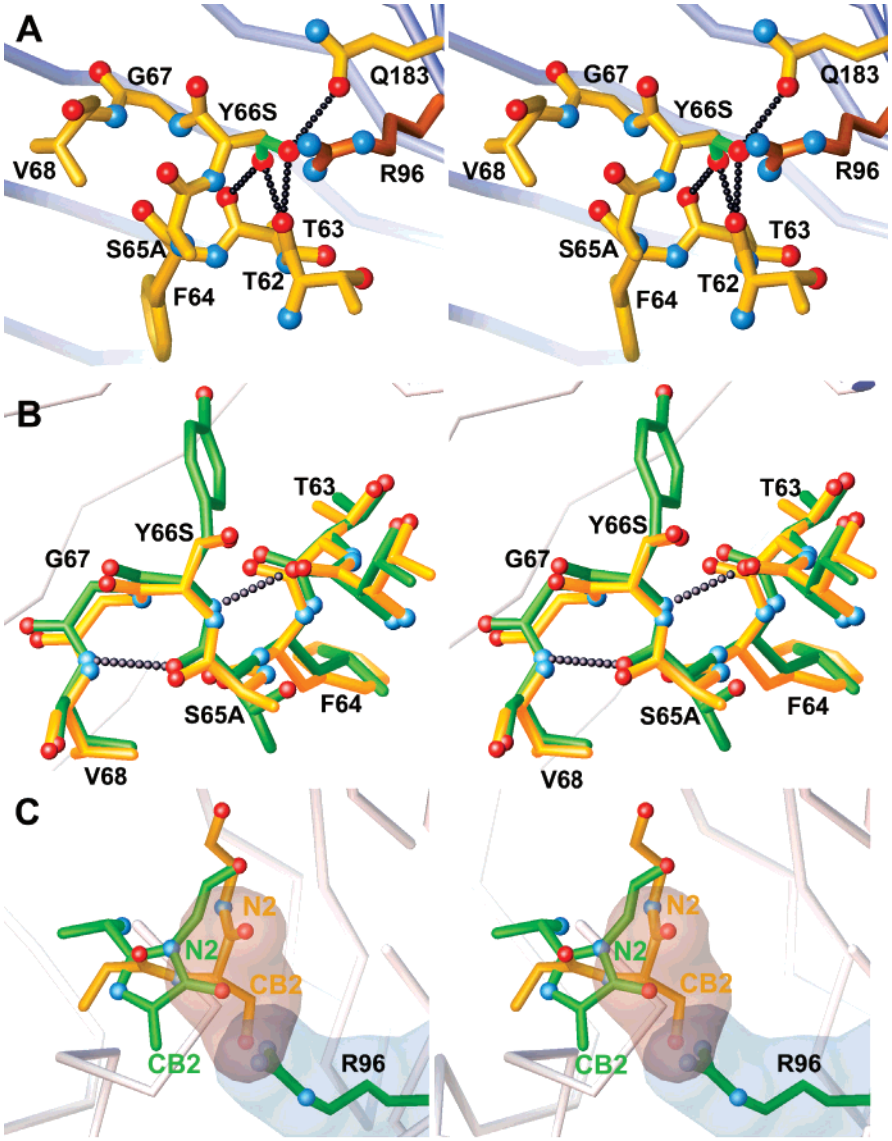

FIGURE 5: Crystallographic structure and analysis of the GFPhal R96A variant. (A) Stereopair of the GFPhal R96A variant (yellow) displayed with the two Y66S conformations (green) and modeled wild-type R96 (brown). (B) Stereopair for structural overlay of GFPhal R96A (yellow) and R96A (green) variants. (C) Stereopair for structural overlay of GFPhal R96A (yellow) and GFPhal (green) structures emphasizing movement induced by R96. van der Waal surface shown for R96 (light blue) and Y66S (light red).

Structure and Characterization of the GFPhal H148G Variant. To explore the potential reactivity of the GFPhal MIO, we next added the H148G side-chain truncation mutation (*39*, *40*), thereby opening solvent/substrate access to the modified chromophore. The GFPhal H148G variant exhibits an absorbance spectrum similar to that of GFPhal but with a lower A385/A280 ratio (Figure 4A), suggesting either a fundamentally distinct chromophore or a different mixture of maturation products. The A385/A280 ratios for each of these two variants remain constant over pH and time (data not shown), indicating that the difference is not due to a slow kinetic maturation event. Interestingly, our 1.50 Å resolution GFPhal H148G structure (Table 1 and Figure 4C) reveals a cyclized product (Figure 1F) that has undergone an Y66S side-chain elimination reaction, similar to the GFPhal variant, but contains a hydroxyl group bound to the C1 atom (the S65A carbonyl carbon atom, Figure 1D), similar to the "dehydration-compromised" Gly-Gly-Gly (see Figure 1C) (*28*) and Y66L (*21*) variants. Outside of the altered chromophore and the H148G mutation, the structures (Figure 4D) are highly similar (overall Cα rmsd of 0.1 Å). Occupancy refinement of the bound hydroxyl group for GFPhal H148G indicates a population greater than 85%, while careful inspection of the omit electron density maps for the GFPhal variant (above) indicated the presence of a small population (estimate of <10%) of bound hydroxyl group. Thus, the GFPhal variant structures reveal a mixture of aromatic and nonaromatic species with the GFPhal variant mostly aromatic (Figure 1E) and the GFPhal H148G variant (Figure 1F) largely nonaromatic.

To test for potential GFPhal MIO reactivity, similar to the modified/inhibited MIOs observed in HAL (*30*) and PAL (*8*), we initiated solution and crystallographic soaking experiments with exogenous nucleophiles and observed no indication of a modified MIO moiety. The apparent lack of reactivity is not due to limited nucleophile access, because the opening created by the H148G truncation allowed a cryo-solvent ethylene glycol molecule (Figure 4C) to approach and form a hydrogen bond to the N2 atom of the MIO. Alternatively, structural studies on the inactive HAL variants

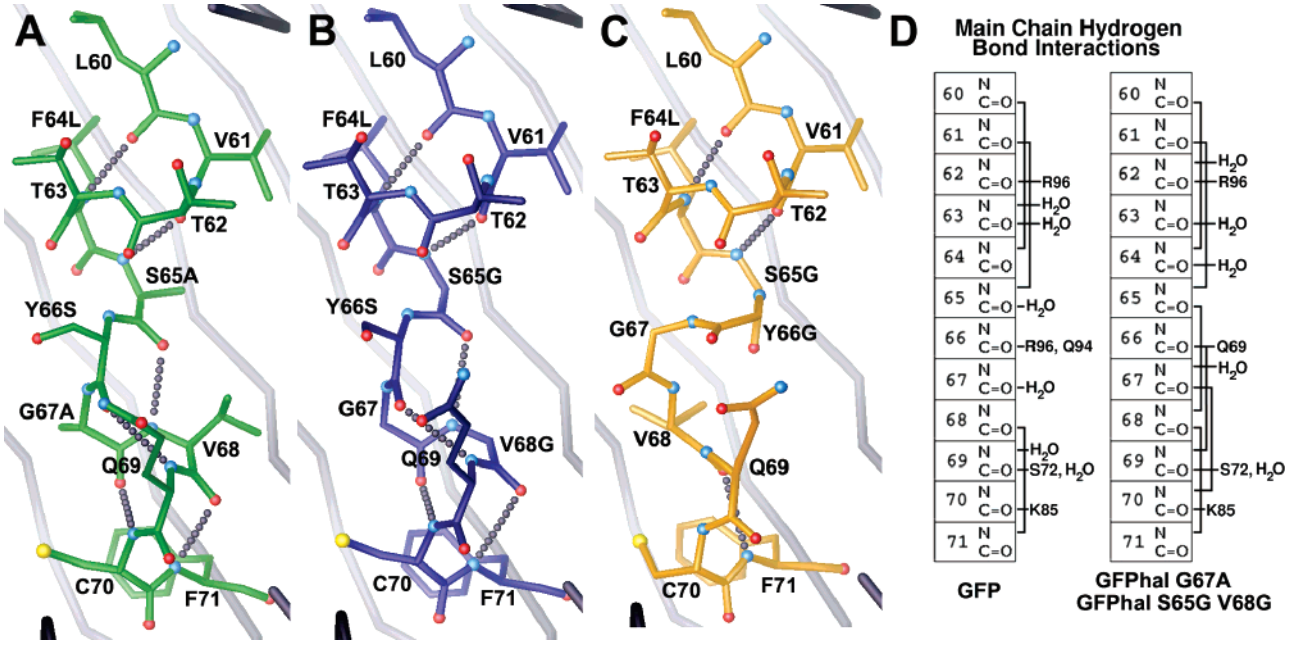

FIGURE 6: Crystallographic structures and analysis of the conformational switch imposed by the GFPhal G67A and GFPhal S65G V68G variants. Conformation for the central helix with hydrogen-bond interactions displayed for the (A) GFPhal G67A, (B) GFPhal S65G V68G, and (C) Gly-Gly-Gly precursor variant structures. D) Schematic for the main-chain hydrogen-bond interactions of wild-type (and Gly-Gly-Gly) structures compared to those of GFPhal G67A and GFPhal S65G V68G. There are no significant conformational differences for the two subunits of the GFPhal G67A variant (overall Cα rmsd of 0.26 Å).

Y280F and F329A indicate that MIO reactivity may critically depend on the sp^3 hybridization of the N3 atom (*20*, *30*). In GFP, R96 has been proposed to facilitate formation of a sp^2-hybridized N3 atom through electrostatic destabilization and deprotonation (*28*). Thus, to potentially stabilize a HAL-like sp^3-hybridized and reactive MIO moiety and further mimic HAL electrophile formation, we constructed and characterized the GFPhal R96A variant.

Structure and Characterization of the GFPhal R96A Variant. The 1.90 Å resolution structure (Table 1) of the GFPhal R96A variant reveals a noncyclized conformation (Figure 5A) and supports a role for R96 in organizing the GFP central helix for backbone cyclization. Previously, the rate of chromophore formation for the R96A variant was shown to be greatly attenuated compared to wild-type GFP, occurring on the time frame of weeks to months (*28*). The GFPhal R96A variant appears to be even slower, with no visible absorbance or structural indication of MIO formation after room-temperature incubation for many months. The noncyclized structure has an intact Y66S side chain, which indicates that backbone cyclization and/or the R96 side chain are important for the hydroxyl elimination reaction observed in GFPhal and GFPhal H148G. In GFPhal R96A, the Y66S side chain has dual conformations that differ by an ~120° chi^{-1} torsion, allowing the side-chain hydroxyl to form hydrogen bonds with the backbone carbonyls of T62 and T63 in one conformation and to interact with the T62 carbonyl and Q183 side-chain Oϵ1 atoms in a second conformation (Figure 5A). The central helix (residues 60−71) conformation of the GFPhal R96A is similar (Figure 5B) to that of the precyclized R96A structure (*28*), as is the overall structure (Cα rmsd of 0.7 Å). To generate the chromophore, the Y66S side-chain hydrogen bonds must be broken and the backbone atoms for the chromophore-forming residues shifted. Significantly, both the altered backbone conformation and the Y66S conformers would be disfavored by steric overlap with R96 (parts A and C of Figure 5), implicating the R96 side chain in preorganization of the chromophore sequence for cyclization. Because GFPhal R96A did not form an MIO, we could not probe the role of R96 in electrostatically destabilizing a positively charged sp^3-hybridized N3 atom. Hence, we generated the GFPhal G67A and GFPhal S65G V68G mutants intended to better mimic HAL by altering the environment of the N3 atom to favor its sp^3 hybridization after cyclization. The GFPhal G67A variant was designed on the rationale that (1) the equivalent mutation in HAL creates a functional MIO species (*30*) and (2) GFP G67 variants are nonfluorescent (*10*), possibly indicating a nonaromatic ring system with a sp^3-hybridized N3 atom. Moreover, the GFPhal S65G V68G variant truncates residues that flank the MIO moiety, thereby relaxing architectural constraints that may be important in favoring the sp^2-hybridized N3 atom and aromatic ring system of the wild-type chromophore.

Structural Rearrangement for the GFPhal G67A and GFPhal S65G V68G Variants. The GFPhal G67A and GFPhal S65G V68G variants undergo a common backbone conformational switch (Figure 6A), and both remain uncyclized. Our 1.50 Å resolution crystal structure (Table 1) of the GFPhal G67A variant reveals a precursor noncyclized conformation for the chromophore atoms, significant localized backbone rearrangements with large (>3.5 Å) Cα movement for the G67A and V68 residues (Figure 6A), and a coil−3_{10}-helix transition for the chromophore-forming residues (S65A−F71). This rearrangement is likely initiated by steric contacts between the G67A Cβ atom and the T63 carbonyl (*28*) and results in placement of the Val68 side chain into a pocket near L42 and E222, displacing water molecules that previously formed hydrogen bonds with the E222 side chain. Interestingly, the 2.5 Å structure of the GFPhal S65G V68G variant (Table 1) reveals a remarkably similar uncyclized conformation (Figure 6B) to that of the

GFPhal G67A variant. The Y66S side chain forms hydrogen bonds to R96 and Q94, which, in wild-type GFP, both form hydrogen bonds with the Y66 carbonyl oxygen. The same coil−3_{10}-helix transition observed in the GFPhal G67A variant is induced in the GFPhal S65G V68G variant.

The new 3_{10} helix stabilizes the precursor state and inhibits chromophore biosynthesis in these noncyclizing GFPhal variants. The number of main-chain helical hydrogen bonds in the distorted precyclized central helix of GFP doubles from 3 to 6 (Figure 6). The three new main-chain hydrogen bonds (to the carbonyl oxygen atoms of the uncyclized chromophore tripeptide, Figure 6D) must be broken to complete chromophore maturation. In these GFPhal variants, the short (~3 Å) distance from the nitrogen nucleophile to its carbonyl carbon target matches that of the Gly-Gly-Gly mutant before cyclization (*28*), but the 3_{10}-helical transition rearranges residue 67 so that its nitrogen lone pair is no longer oriented for nucleophilic attack. These results suggest that 3_{10}-helix formation, triggered by increased (GFPhal G67A) or decreased (GFPhal S65G V68G) steric constraints, inhibits tripeptide cyclization by preferential stabilization of the precursor state.

DISCUSSION

Our approach to understanding the driving force and post-translational chemistry for self-catalyzed reactions within proteins is to generate specific structure-based hypotheses and test these ideas with site-directed mutagenesis and high-resolution crystallography. For example, in the HAL/PAL protein architecture, seven helix dipoles have been proposed to increase the electrophilicity and reactivity of the MIO moiety (*8*). Although difficult to test within native protein scaffolds, such an architecture-specific proposal can be probed by transplanting the MIO into a different target scaffold. Toward that end, we transplanted the MIO electrophile tripeptide sequence into the GFP scaffold, altered the surrounding electrostatic environment through site-directed mutagenesis, and evaluated the MIO reactivity. In GFP, the transplanted HAL tripeptide sequence indeed cyclizes but produces an aromatic MIO that differs from the nonaromatic MIO in HAL by hybridization differences at the N3 atom. In wild-type HAL, sp^3-hybridization geometry at the N3 atom, defined by the tetrahedral geometry of its covalently bound atoms, excludes the nitrogen lone pair from conjugation with the rest of the imidazolone ring. Thus, the HAL electrophile is nonaromatic and does not exhibit a visible absorbance spectrum. In contrast, the GFPhal variant and the low activity HAL mutants Y280F and F329A (*20*, *30*) have sp^2-hybridization geometry at the N3 atom, allowing conjugation of its lone pair electrons. The observed differences in N3 hybridization and consequent MIO aromaticity in GFPhal but not in HAL arise from distinct biosynthetic features of the two scaffolds and produce important functional differences.

Driving Force for Peptide Backbone Cyclization. Both the GFP and HAL protein frameworks locally destabilize the conformation of the tripeptide precursor to favor cyclization. In GFP, probably to inhibit fluorescence quenching, the reaction occurs on a distorted α helix deep inside the β barrel, whereas the HAL reaction occurs on a solvent-exposed loop that allows substrate access. Cyclization is a backbone condensation reaction that would be inhibited by main-chain hydrogen bonds present in the precursor state but not in the cyclized product. In GFP, the central chromophore-containing helix exhibits a dramatic bend centered at the chromophore that precludes all but 6 of 24 possible main-chain hydrogen-bonding interactions in both precyclization and mature states (Figure 6D). These distortions specifically align molecular orbitals for bond formation (see below) and remove inhibitory hydrogen bonds with the G67 nucleophile, which forms the new covalent bond, and the S65 carbonyl oxygen atom, which is lost as a water molecule (*28*). Because the central helix distortions are present in both precyclized and mature structures, the conjugation-trapping model suggests that these distortions do not drive GFP peptide backbone cyclization per se but allow population of an unstable cyclized intermediate, which is then trapped through a subsequent reaction (*28*). In contrast, HAL drives the cyclization reaction through an apparent mechanical compression mechanism (*20*), in which steric constraints imposed by the protein scaffold are relaxed in forming the MIO moiety, implying that the initial cyclized product, unlike GFP, is more stable than the precursor.

Negative Design. Fluorophore biosynthesis by GFP has evolved to incorporate principles of both positive (favoring the desired conformation) and negative (disfavoring alternative undesirable conformations) protein design. Such principles underscore the importance of both favoring functional and eliminating nonfunctional options for biological systems (*41*, *42*). As described above, the wild-type GFP architecture incorporates positive design features that lower energetic barriers to ring closure by precluding inhibitory hydrogen bonds and by imposing a specific conformation on the chromophore tripeptide. In contrast, the GFP precursor structures (parts A−C of Figure 6) highlight specific negative design features present in wild-type GFP but missing for these GFP variants, which are defective in backbone cyclization and chromophore maturation. For the noncyclized GFPhal R96A variant, unlike the wild type, the precyclization state is preferentially stabilized through Y66S hydrogen-bond interactions and backbone movements that raise the energetic cost of peptide cyclization. Compared to wild-type GFP, the GFPhal G67A and GFPhal S65G V68G variants form three additional 3_{10}-helical hydrogen bonds (65−68, 66−69, and 67−70), which opposes or must be broken for backbone cyclization. Thus, a common feature for these stalled chromophore maturation variants is the preferential stabilization of the precursor state, suggesting that negative design features of wild-type GFP destabilize these nonproductive conformations to favor chromophore formation.

Precursor Architecture for Peptide Backbone Cyclization. GFP favors chromophore biosynthesis through a specific architecture that aligns molecular orbitals for bond formation. Previously, we showed (*28*) that (1) the Gly-Gly-Gly GFP variant undergoes backbone cyclization but, in the absence of oxygen, reverts back to the precyclization state and (2) entropic barriers are reduced by the close van der Waals nitrogen−carbon contact distance (3.0 Å) for the ring-forming atoms. Further analysis of the Gly-Gly-Gly precyclization structure (PDB code 1QYO) reveals that the 65−66 peptide bond is oriented perpendicular to that of 66−67. Importantly, this arrangement aligns the G67 nitrogen lone pair with an empty π^* antibonding orbital of the carbonyl

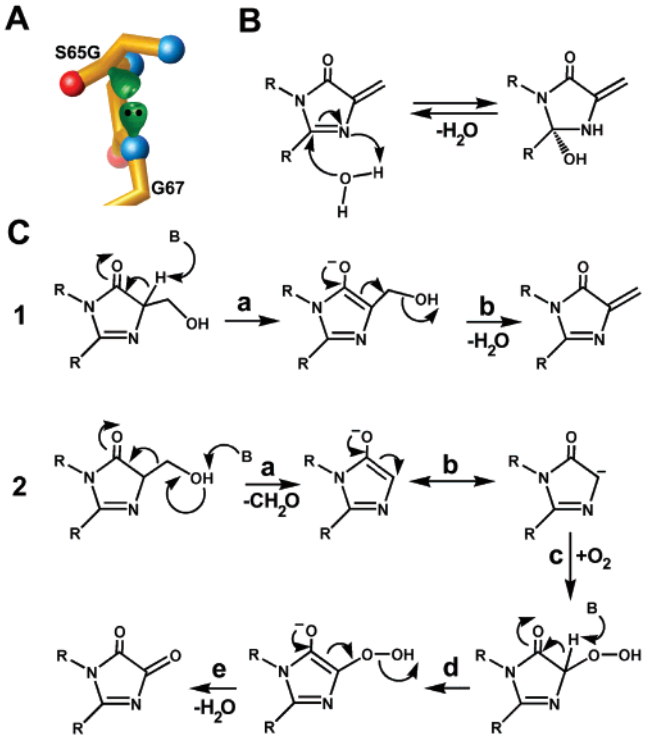

FIGURE 7: Post-translational chemistry for GFP variants. (A) Orbital alignment for peptide backbone cyclization for the viable Gly-Gly-Gly precyclized (anaerobic) structure. The lone pair and π^* orbitals are displayed in green. (B) Reversible dehydration/hydration reactions for the GFPhal variants. (C) Post-translational products generated through HAL-like dehydration (scheme 1) and formaldehyde elimination (scheme 2) reactions.

(Figure 7A). Although the GFPhal G67A and GFPhal S65G V68G variants exhibit similar short nitrogen to carbonyl carbon contacts (3.0−3.3 Å) for the bond-forming atoms, the coil−helix transition results in substantially different and nonoptimized nitrogen−carbonyl orbital alignments. Further, comparison of the GFPhal and GFPhal R96A structures indicates that R96 induces structural rearrangements (Figure 5C) that align orbitals for bond formation and may help explain the inhibited cyclization for the R96A and GFPhal R96A variants.

Reversible Main-Chain Dehydration Reaction. The cyclized GFPhal variants reveal a mixture of aromatic (dehydrated) and nonaromatic (nondehydrated) moieties (Figures 1 and 4). The GFPhal and GFPhal H148G variant structures exhibit no electron density for the Y66S side-chain hydroxyl, yet they have different populations for the S65A main-chain carbonyl oxygen. Structural overlay between the GFPhal and GFPhal H148G variants reveals minor changes between the dehydrated (with proximal water Wat2) and nondehydrated moieties (Figure 4D). We establish a correlation between the spectroscopic and structural properties for the GFPhal variants that suggests a mixed population of aromatic and nonaromatic species that does not change over time. These data are consistent with two energetically similar thermodynamic products and a reversible dehydration/hydration reaction (Figure 7B). A similar reversible hydration reaction was postulated for wild-type GFP between a proposed enamine intermediate and the mature chromophore (*21*), and an energetically similar aromatic/nonaromatic transition was hypothesized for intermediates in HAL catalysis (*30*).

A reversible dehydration/hydration reaction has implications for the interpretation of previous nondehydrated GFP variants. First, the cyclization−oxidation−dehydration mechanism is based on the presence of a nondehydrated main-chain hydroxyl in the Y66L structure (*21*). Our results for the GFPhal variants suggest that the hydration/dehydration of the imidazolone ring is reversible. Thus, the presence of the hydroxyl group in the Y66L structure, especially without data indicating that this species is an intermediate in chromophore maturation, cannot be extrapolated to define the kinetic mechanism (see below). Currently, the order of the main-chain dehydration and oxidation (in GFP) or side-chain dehydration (in HAL/GFPhal) steps remains unresolved. Reported (unpublished) mass spectrometry experiments in GFP suggest that the mass difference between the anaerobic and mature chromophore states is consistent with oxidation but not dehydration (*10*). However, these mass spectrometry results cannot rule out a reversible equilibrium between a more stable dehydrated moiety and a low population of a hydroxide-bound species that undergoes oxidation to form the mature chromophore.

Second, both the orbital alignment for bond formation in the Gly-Gly-Gly precyclized structure and the reversible hydration/dehydration equilibrium in GFPhal variants are consistent with deprotonated G67 nitrogen nucleophilic attack on the S65 carbonyl. The superimposable hydroxylated and dehydrated GFPhal structures reported here (Figure 4D) lead to the conclusion that dehydration is reversible. Thus, our previous assumption that the Gly-Gly-Gly variant is "dehydration-compromised" because of a shift of the imidazolone ring and the resultant proposal for G67 deprotonation after ring cyclization (*28*) are not needed. These new data, coupled to the alignment of the G67 nitrogen lone pair with the empty π^* antibonding orbital of the carbonyl (Figure 7A), indicate that G67 is instead deprotonated prior to ring cyclization (a proton on the G67 nitrogen would inhibit ring formation), supporting previous hypotheses (*18*, *21*, *22*). G67 deprotonation may be assisted by ionic interactions between the R96 guanidinium side chain and the Y66 carbonyl, which favor the peptide bond resonance form with double-bond character (*28*).

GFPhal Post-translational Modification. The crystallographic structures of GFPhal and GFPhal H148G reveal post-translational modifications with eliminated Y66S side-chain hydroxyl groups. These data indicate that, unlike wild-type GFP, GFPhal has *not* undergone a Cα−Cβ oxidation reaction, which would result in vicinal alcohol or aldehyde moiety. Throughout the results section, we assumed a HAL-like side-chain dehydration mechanism, in which the Y66S hydroxyl group is lost as water, leaving behind a double bond (between CA2 and CB2) exocyclic to the five-membered ring (scheme 1 of Figure 7C). Interestingly, the GFPhal structure reveals a very short distance (2.27 ± 0.07 Å) between the CB2 atom and the oxygen atom of a water molecule (Wat1, Figure 4B), indicating an unusual carbon−oxygen hydrogen bond, likely favored by the increased acidity of the CB2 methylene carbon of the imidazolone ring. Alternatively, a more complicated formaldehyde elimination and oxygen incorporation mechanism (scheme 2 of Figure 7C) would result in a species with two exocyclic carbonyl substituents similar to that previously proposed for the Gly-Gly-Gly variant (*28*). In this case, a conventional oxygen−

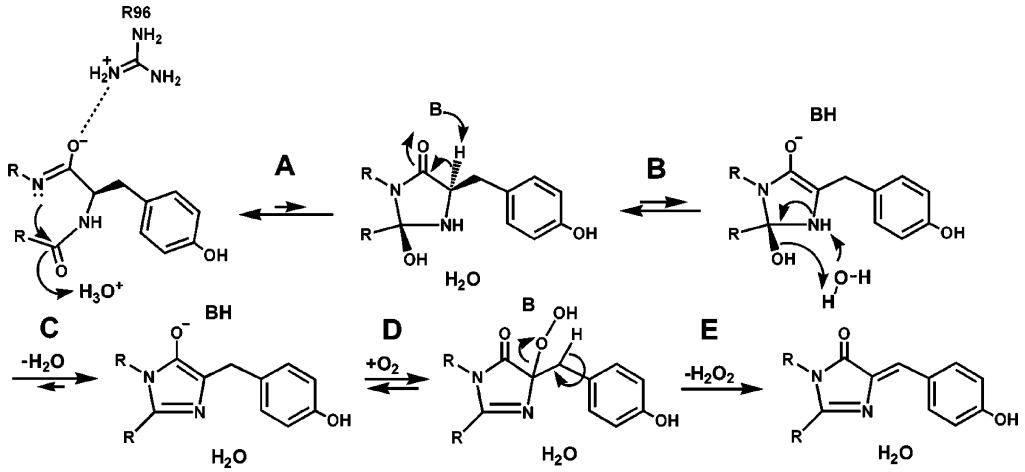

FIGURE 8: Proposed biosynthetic scheme for GFP chromophore formation. (A) Thermodynamically unfavorable backbone cyclization reaction. (B) CA2 deprotonation and enolate formation. R96 electrostatic interactions would stabilize this intermediate. (C) Reversible main-chain dehydration reaction. This aromatic intermediate would be further stabilized by R96. (D) Oxygen addition and hydroperoxide adduct formation. (E) Oxidation through CB2 deprotonation and hydrogen peroxide generation.

water hydrogen bond would be formed between the newly incorporated oxygen atom and Wat1. Our crystallographic data cannot distinguish a modified chromophore containing a methylene group [molecular weight (MW) 14; scheme 1 of Figure 7C] from one with an incorporated oxygen atom (MW 16; scheme 2 of Figure 7C). In the absence of additional spectroscopic or biochemical data, the observed bond lengths lead us to tentatively favor an exocyclic carbon atom and the simpler HAL-like dehydration mechanism (scheme 1 of Figure 7C). If this molecular assignment can be confirmed, these data provide the first evidence that neither GFP cyclization nor stabilization of a GFP-cyclized product depends on an oxidation reaction.

GFP Chromophore Formation Mechanism. The structures of three GFP variants, Gly-Gly-Gly (S65G Y66G) (*28*), Y66L (S65T Y66L) (*21*), and GFPhal (S65A Y66S) (presented here) with modified imidazolone ring systems shed light on the mechanism of GFP chromophore formation. First, the Gly-Gly-Gly variant incorporated a single oxygen atom attached to the Y66G Cα position and contained a hydroxyl group bound to the S65G carbonyl carbon, suggesting that this variant was dehydration-compromised (*28*). Under anaerobic but not aerobic conditions, the Gly-Gly-Gly variant exhibits a precyclized conformation, suggesting two possibilities: oxidation is required prior to ring cyclization or oxygen is necessary to stabilize the cyclized product. On the basis of these results, we proposed the conjugation-trapping model for GFP chromophore formation, in which a thermodynamically unfavorable cyclization reaction is trapped by a subsequent stabilizing reaction, such as oxidation for the Gly-Gly-Gly variant or dehydration for wild-type GFP (*28*).

Second, Wachter and colleagues determined the Y66L crystallographic structure and explained the observed sp^2-hybridized Y66L Cα atom of the cyclized and presumably oxidized chromophore with three possible "intermediate" structures: an enolate (CA2−C2 double bond), an oxidized heterocyclic moiety (CA2−N2 double bond), or an enamine tautomer of that oxidized heterocyclic species (CA2−CB2 double bond) (*21*). In Y66L, as in the Gly-Gly-Gly variant, the modified chromophore had a hydroxyl group attached to the residue 65 carbonyl carbon atom. Rosenow et al. proposed a cyclization−oxidation−dehydration mechanism for wild-type GFP, incorporating hydroperoxide adduct formation with the CA2 atom, N2 deprotonation with hydrogen peroxide generation, and CB2 proton abstraction to eject the water molecule. Further, these authors concluded that the substitution of the aromatic Y66 with the aliphatic Leu inhibits dehydration by impairing CB2 proton abstraction.

Third, GFPhal variant spectroscopic and structural data reveal mixtures of aromatic (dehydrated) and nonaromatic (nondehydrated) products (parts E and F of Figure 1) for the same completed modification at Y66 Cα (side-chain elimination producing a dehydroalanine with a CA2−CB2 double bond). Thus, modifications at the Y66 Cα position are not necessarily coupled to the main-chain dehydration reaction. To explain the mixed population of dehydrated species, we proposed a reversible N2 deprotonation and water elimination reaction (Figure 7B), similar to that proposed by Cubitt et al. (*18*) for wild-type GFP. In contrast, the Rosenow et al. model requires two assumptions to explain the GFPhal dehydrated product: N2−CA2 tautomerization for the GFPhal imidazolone ring and much greater CB2 acidity for the GFPhal dehydroalanine methyl substituent than for the branched aliphatic Leu side chain of Y66L.

Here, we modify and extend previous chromophore formation schemes to accommodate all of the current data. First, we agree with the generally accepted (but not proven) proposal by Heim et al. (*14*) that the first step in chromophore formation is ring cyclization (Figure 8A). On the basis of the alignment of the G67 nucleophile with the residue 65 carbonyl in the Gly-Gly-Gly precursor structure (Figure 7A), the G67 nitrogen appears to be deprotonated prior to ring formation. Second, after ring cyclization, we propose that proton abstraction from CA2 facilitates enolate formation (Figure 8B). At this point, there are two primary options: N2 deprotonation and dehydration, originally proposed by Cubitt et al. (*18*), or hydroperoxide formation at the CA2 atom, as postulated for the Gly-Gly-Gly (*28*) and Y66L (*21*) variants. Currently, no published data distinguish these possibilities (see above). Third, we favor the Cubitt et al. dehydration reaction prior to oxidation (Figure 8C). The resultant aromatic ring system helps to stabilize the product

of the electronically unfavorable cyclization reaction (*20*) while awaiting the slow [$t_{1/2} \sim 4$ h (*14*)] oxidation reaction, which likely requires molecular oxygen. R96 electrostatic interactions would further stabilize this enolate intermediate and help explain the absolute conservation of R96 in GFP homologues. Fourth, molecular oxygen incorporation generates a hydroperoxide adduct (Figure 8D). The proposed CA2 adduct is consistent with the incorporation of an exocyclic oxygen atom in the Gly-Gly-Gly variant (*28*). Fifth, CB2 deprotonation and hydrogen peroxide loss leads to the generation of the mature fluorophore (Figure 8E). We favor loss of the hydroperoxide adduct through CB2 deprotonation, rather than N2 deprotonation (*21*), to better explain the Gly-Gly-Gly structural data (*28*); the lack of a CB2 atom leads to CA2 proton abstraction, water elimination, and single oxygen incorporation at the Y66G Cα atom (Figure 1C). Importantly, if these modified chromophore results and hypotheses extend to and provide insights for the biosynthetic mechanism of wild-type GFP, they may lead to the design of variants with faster chromophore maturation kinetics or desired photophysical properties.

Summary. Structure-guided protein engineering offers opportunities to define the determinants for post-translational modifications and to build proteins with novel properties. Design is the ultimate test of understanding. Here, we successfully re-engineer the chemistry for the GFP post-translational modification and convert the fluorophore into an electrophilic moiety, similar to the active-site catalyst of histidine ammonia lyase. Our crystallographic structures and analysis reveal negative design principles that underscore the importance of both favoring functional and eliminating nonfunctional options for biological systems. Significantly, our results and analysis facilitate the in-depth understanding of the driving force and protein chemistry that control backbone cyclization and chromophore creation. The protein architecture promotes post-translational backbone cyclization by (1) favoring nucleophilic attack by close proximity alignment of the G67 amide lone pair with the π^* orbital of the residue 65 carbonyl, and (2) removing enthalpic barriers by eliminating inhibitory main-chain hydrogen bonds in the precursor state. These requirements are not necessarily scaffold-specific and therefore suggest opportunities to engineer modified amino acids into target scaffolds.

ACKNOWLEDGMENT

We thank C. D. Putnam, V. A. Roberts, T. I. Wood, J. L. Huffman, and C. A. Mullen for scientific discussions and helpful suggestions and K. D. Mullen for assistance during the preparation of this manuscript.

REFERENCES

1. Okeley, N. M., and van der Donk, W. A. (2000) Novel cofactors via post-translational modification of enzyme active sites, *Chem. Biol. 7*, R159−R171.
2. Ormö, M., Cubitt, A. B., Kallio, K., Gross, L. A., Tsien, R. Y., and Remington, S. J. (1996) Crystal structure of the *Aequorea victoria* green fluorescent protein, *Science 273*, 1392−1395.
3. Yang, F., Moss, L. G., and Phillips, G. N., Jr. (1996) The molecular structure of green fluorescent protein, *Nat. Biotech. 14*, 1246−1251.
4. Poppe, L. (2001) Methylidene-imidazolone: A novel electrophile for substrate activation, *Curr. Opin. Chem. Biol. 5*, 512−524.
5. Retey, J. (2003) Discovery and role of methylidene imidazolone, a highly electrophilic prosthetic group, *Biochim. Biophys. Acta 1647*, 179−184.
6. Langer, B., Langer, M., and Retey, J. (2001) Methylidene-imidazolone (MIO) from histidine and phenylalanine ammonia-lyase, *Adv. Protein Chem. 58*, 175−214.
7. Schwede, T. F., Retey, J., and Schulz, G. E. (1999) Crystal structure of histidine ammonia-lyase revealing a novel polypeptide modification as the catalytic electrophile, *Biochemistry 38*, 5355−5361.
8. Calabrese, J. C., Jordan, D. B., Boodhoo, A., Sariaslani, S., and Vannelli, T. (2004) Crystal structure of phenylalanine ammonia lyase: Multiple helix dipoles implicated in catalysis, *Biochemistry 43*, 11403−11416.
9. Christenson, S. D., Liu, W., Toney, M. D., and Shen, B. (2003) A novel 4-methylideneimidazole-5-one-containing tyrosine aminomutase in enediyne antitumor antibiotic C-1027 biosynthesis, *J. Am. Chem. Soc. 125*, 6062−6063.
10. Tsien, R. Y. (1998) The green fluorescent protein, *Annu. Rev. Biochem. 67*, 509−544.
11. Zimmer, M. (2002) Green fluorescent protein (GFP): Applications, structure, and related photophysical behavior, *Chem. Rev. 102*, 759−782.
12. Matz, M. V., Fradkov, A. F., Labas, Y. A., Savitsky, A. P., Zaraisky, A. G., Markelov, M. L., and Lukyanov, S. A. (1999) Fluorescent protein from nonbioluminescent anthozoa species, *Nat. Biotech. 17*, 969−973.
13. 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.
14. Heim, R., Prasher, D. C., and Tsien, R. Y. (1994) Wavelength mutations and posttranslational autooxidation of the green fluorescent protein, *Proc. Natl. Acad. Sci. U.S.A. 91*, 12501−12504.
15. Heim, R., and Tsien, R. Y. (1996) Engineering green fluorescent protein for improved brightness, longer wavelengths, and fluorescence resonance energy transfer, *Curr. Biol. 6*, 178−182.
16. Taylor, R. G., Levy, H. L., and McInnes, R. R. (1991) Histidase and histidinemia. Clinical and molecular considerations, *Mol. Biol. Med. 8*, 101−116.
17. Sarkissian, C. N., Shao, Z., Blain, F., Peevers, R., Su, H., Heft, R., Chang, T. M., and Scriver, C. R. (1999) A different approach to treatment of phenylketonuria: Phenylalanine degradation with recombinant phenylalanine ammonia lyase, *Proc. Natl. Acad. Sci. U.S.A. 96*, 2339−2344.
18. Cubitt, A. B., Heim, R., Adams, S. R., Boyd, A. E., Gross, L. A., and Tsien, R. Y. (1995) Understanding, improving, and using green fluorescent proteins, *Trends Biochem. Sci. 20*, 448−455.
19. Reid, B. G., and Flynn, G. C. (1997) Chromophore formation in green fluorescent protein, *Biochemistry 36*, 6786−6791.
20. Baedeker, M., and Schulz, G. E. (2002) Autocatalytic peptide cyclization during chain folding of histidine ammonia-lyase, *Structure 10*, 61−67.
21. Rosenow, M. A., Huffman, H. A., Phail, M. E., and Wachter, R. M. (2004) The crystal structure of the Y66L variant of green fluorescent protein supports a cyclization−oxidation−dehydration mechanism for chromophore maturation, *Biochemistry 43*, 4464−4472.
22. Siegbahn, P. E. M., Wirstam, M., and Zimmer, M. (2001) Theoretical study of the mechanism of peptide ring formation in green fluorescent protein, *Int. J. Quantum Chem. 81*, 169−186.
23. Gross, L. A., Baird, G. S., Hoffman, R. C., Baldridge, K. K., and Tsien, R. Y. (2000) The structure of the chromophore within DsRed, a red fluorescent protein from coral, *Proc. Natl. Acad. Sci. U.S.A. 97*, 11990−11995.
24. Wall, M. A., Socolich, M., and Ranganathan, R. (2000) The structural basis for red fluorescence in the tetrameric GFP homolog DsRed, *Nature Struct. Biol. 7*, 1133−1138.
25. 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.
26. Branchini, B. R., Nemser, A. R., and Zimmer, M. (1998) A computational analysis of the unique protein-induced tight turn that results in posttranslational chromophore formation in green fluorescent protein, *J. Am. Chem. Soc 120*, 1−6.
27. Donnelly, M., Fedeles, F., Wirstam, M., Siegbahn, P. E., and Zimmer, M. (2001) Computational analysis of the autocatalytic posttranslational cyclization observed in histidine ammonia-lyase.

A comparison with green fluorescent protein, *J. Am. Chem. Soc. 123*, 4679−4686.
28. Barondeau, D. P., Putnam, C. D., Kassmann, C. J., Tainer, J. A., and Getzoff, E. D. (2003) Mechanism and energetics of green fluorescent protein chromophore synthesis revealed by trapped intermediate structures, *Proc. Natl. Acad. Sci. U.S.A. 100*, 12111−12116.
29. Heim, R., Cubitt, A. B., and Tsien, R. Y. (1995) Improved green fluorescence, *Nature 373*, 663−664.
30. Baedeker, M., and Schulz, G. E. (2002) Structures of two histidine ammonia-lyase modifications and implications for the catalytic mechanism, *Eur. J. Biochem. 269*, 1790−1797.
31. Barondeau, D. P., Kassmann, C. J., Tainer, J. A., and Getzoff, E. D. (2002) Structural chemistry of a green fluorescent protein Zn biosensor, *J. Am. Chem. Soc. 124*, 3522.
32. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode, *Macromol. Crystallogr., Part A 276*, 307−326.
33. Navaza, J. (1994) AMoRe: An automated package for molecular replacement, *Acta Crystallogr., Sect. A 50*, 157−163.
34. McRee, D. E. (1999) XtalView/Xfit: A versatile program for manipulating atomic coordinates and electron density, *J. Struct. Biol. 125*, 156−165.
35. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, N., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallographic and NMR system (CNS): A new software system for macromolecular structure determination, *Acta Crystallogr., Sect. D 54*, 905−921.
36. Sheldrick, G. M. (1997) High-resolution structure refinement, in *Crystallographic Computing*, Oxford University Press, New York.
37. Brunger, A. T. (1992) The free *R* value: A novel statistical quantity for assessing the accuracy of crystal structures, *Nature 355*, 472−474.
38. Bruns, C. M., Hubatsch, I., Ridderstrom, M., Mannervik, B., and Tainer, J. A. (1999) Human glutathione transferase A4-4 crystal structures and mutagenesis reveal the basis of high catalytic efficiency with toxic lipid peroxidation products, *J. Mol. Biol. 288*, 427−439.
39. Wachter, R. M., Elsliger, M., Kallio, K., Hanson, G. T., and Remington, S. J. (1998) Structural basis of spectral shifts in the yellow-emission variants of green fluorescent protein, *Structure 6*, 1267−1277.
40. Seifert, M. H., Georgescu, J., Ksiazek, D., Smialowski, P., Rehm, T., Steipe, B., and Holak, T. A. (2003) Backbone dynamics of green fluorescent protein and the effect of histidine 148 substitution, *Biochemistry 42*, 2500−2512.
41. Richardson, J. S., Richardson, D. C., Tweedy, N. B., Gernert, K. M., Quinn, T. P., Hecht, M. H., Erickson, B. W., Yan, Y., McClain, R. D., Donlan, M. E. et al. (1992) Looking at proteins: Representations, folding, packing, and design. Biophysical Society National Lecture, *Biophys. J. 63*, 1185−1209.
42. Hellinga, H. W. (1997) Rational protein design: Combining theory and experiment, *Proc. Natl. Acad. Sci. U.S.A. 94*, 10015−10017.

BI0479205