A conserved interaction with the chromophore of fluorescent proteins

Amit Choudhary,[1] Kimberli J. Kamer,[2] and Ronald T. Raines[2,3]*

[1]Graduate Program in Biophysics, University of Wisconsin–Madison, Madison, Wisconsin 53706
[2]Department of Biochemistry, University of Wisconsin–Madison, Madison, Wisconsin 53706
[3]Department of Chemistry, University of Wisconsin–Madison, Madison, Wisconsin 53706

Received 23 August 2011; Revised 25 October 2011; Accepted 28 October 2011
DOI: 10.1002/pro.762
Published online 4 November 2011 proteinscience.org

Abstract: The chromophore of fluorescent proteins, including the green fluorescent protein (GFP), contains a highly conjugated imidazolidinone ring. In many fluorescent proteins, the carbonyl group of the imidazolidinone ring engages in a hydrogen bond with the side chain of an arginine residue. Prior studies have indicated that such an electrophilic carbonyl group in a protein often accepts electron density from a main-chain oxygen. A survey of high-resolution structures of fluorescent proteins indicates that electron lone pairs of a main-chain oxygen—Thr62 in GFP—donate electron density into an antibonding orbital of the imidazolidinone carbonyl group. This n→π* electron delocalization prevents structural distortion during chromophore excitation that could otherwise lead to fluorescence quenching. In addition, this interaction is present in on-pathway intermediates leading to the chromophore, and thus could direct its biogenesis. Accordingly, this n→π* interaction merits inclusion in computational and photophysical analyses of the chromophore, and in speculations about the molecular evolution of fluorescent proteins.

Keywords: green fluorescent protein; hyperconjugation; imidazolidinone; n→π* interaction; stereoelectronic effect

Introduction

Fluorescent proteins, including the green fluorescent protein (GFP) and its analogs, have revolutionized biological imaging.[1–7] At the heart of these fluorescent proteins lies a highly conjugated imidazolidinone ring-based chromophore (Fig. 1). This chromophore appears to have arisen first in an ancient Metazoan, and it has since diverged considerably.[6]

The environment surrounding the chromophore is vital to its spectroscopic properties. For example, the side chain of a proximal arginine residue donates a hydrogen bond to the imidazolidinone carbonyl group. This hydrogen bond should make the imidazolidinone carbonyl group considerably electrophilic. Our prior work on hyperconjugative-type interactions in protein structures suggests that such electrophilic carbonyl groups often engage in a stabilizing interaction with a nucleophile.[8] Such an interaction, which we refer to as an "n→π* interaction,"[9–16] involves delocalization of the electron lone pair (n) of the nucleophilic donor into the antibonding orbital (π*) of the carbonyl group acceptor. This interaction, which is reminiscent of the Bürgi–Dunitz trajectory for nucleophilic addition to a carbonyl group,[17–19] induces a short contact between the nucleophile and the carbonyl carbon in which the van der Waals surfaces of the nucleophile and carbon interpenetrate.[10,12] This interaction has been identified in small molecules, such as γ-aminobutyric acid[20] and aspirin,[21] and larger molecular systems, including peptides,[16] peptoids,[22,23] and proteins,[8] and could have directed the prebiotic genesis of ribonucleotides.[24]

We suspected that the electron-deficient carbonyl carbon of the imidazolidinone ring could interact with a proximal nucleophilic donor. Conversely, we also realized that any such interaction would be

Additional supporting information may be found in the online version of this article.

Grant sponsor: National Institutes of Health; Grant number: R01 AR044276.

*Correspondence to: Ronald T. Raines, Department of Biochemistry, University of Wisconsin–Madison, 433 Babcock Drive, Madison, WI 53706-1544. E-mail: rtraines@wisc.edu

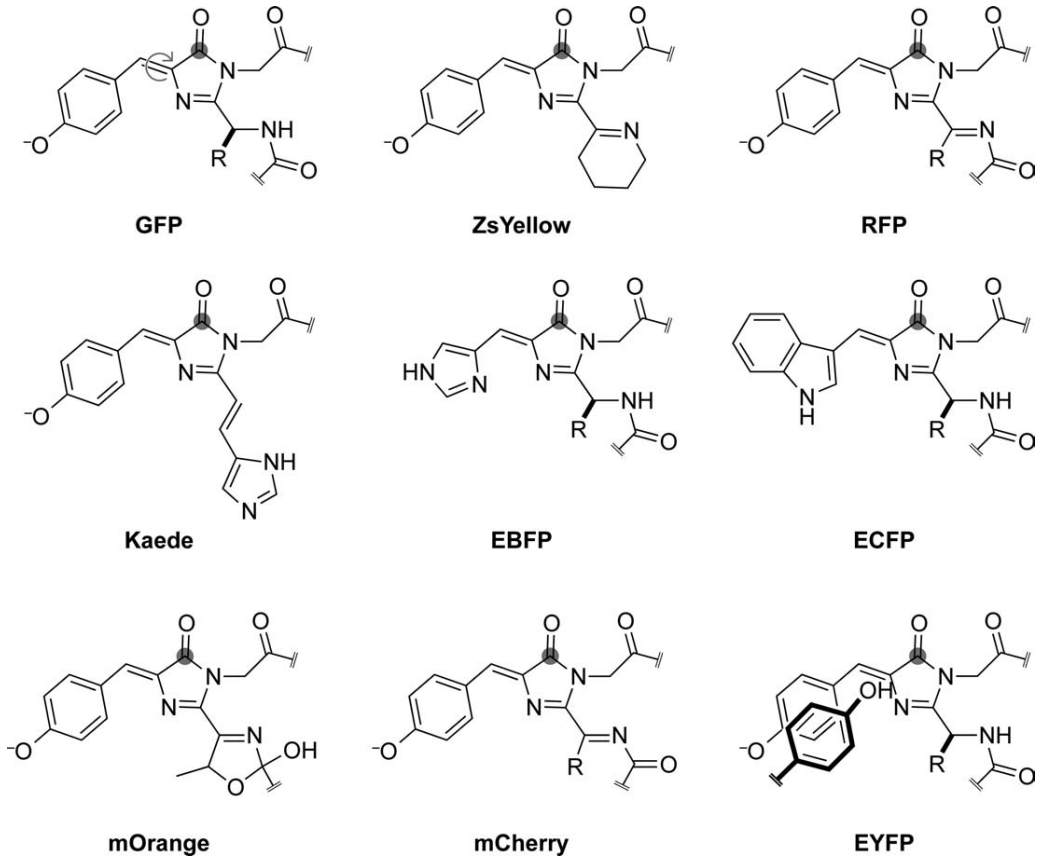

Figure 1. Structure of the chromophores in nine fluorescent proteins. The gray circle indicates the carbon of the carbonyl group that accepts an n→π* interaction. The gray arrow in the GFP chromophore indicates the C—C double bond that can rotate to cause non-radiative decay of the excited state.[38–41]

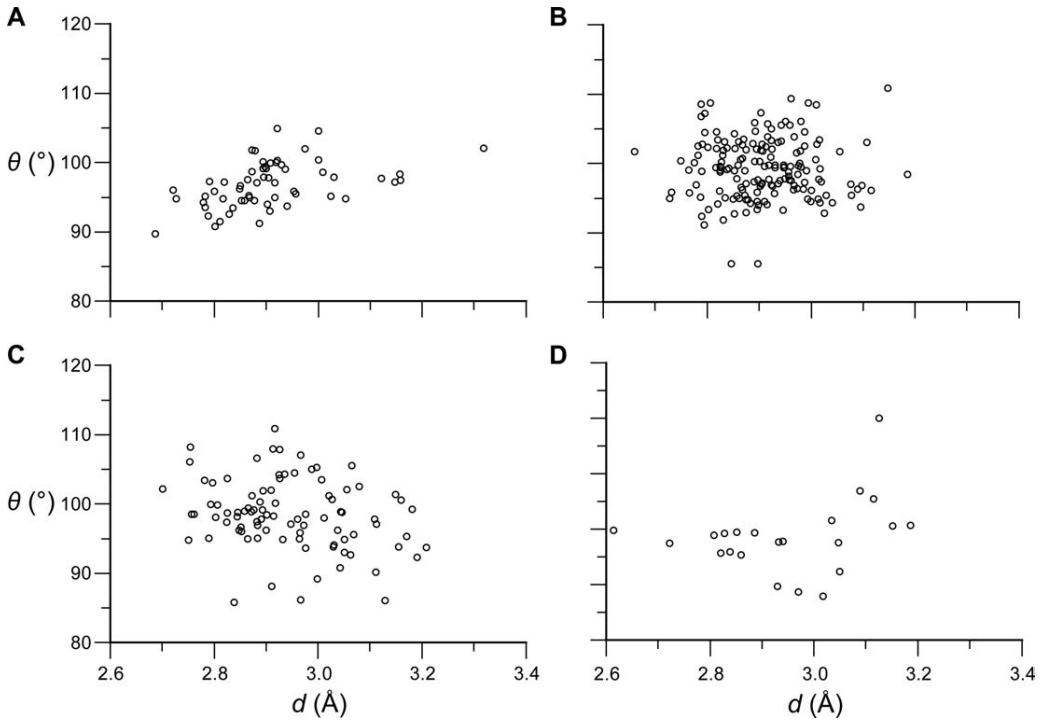

Figure 2. Plot of O···C=O distance (*d*) and angle (θ) for fluorescent protein structures determined at different resolutions: (A) <1.50 Å, (B) 1.50–1.99 Å, (C) 2.00–2.49 Å, and (D) ≥ 2.50 Å.

antagonized by Pauli repulsion between the lone pair of the nucleophilic donor and the π-orbitals of the imidazolidinone ring system.[14] In the context of this landscape, we searched for nucleophilic donors that could donate to the electron-deficient carbonyl carbon in ~200 imidazolidinone-containing fluorescent proteins of known three-dimensional structure in the Protein Data Bank (PDB).[25–27]

Results and Discussion

To our surprise, we found that a main-chain oxygen forms a short contact with the imidazolidinone carbonyl group in each fluorescent protein. Specifically, the distance (d) between the oxygen of the donor carbonyl group and the carbon of the imidazolidinone carbonyl group was less than the sum of the van der Waals radii of oxygen and carbon (3.22 Å). Upon examining the geometry of the n→π* interaction in four subsets of structures determined at comparable resolutions (see: Supporting Information Table S1), we found that the O···C=O angle (θ) formed by the donor oxygen and the acceptor carbonyl group is obtuse and aligned with the Bürgi–Dunitz trajectory (Fig. 2).[12] This interaction between the main-chain oxygen and the imidazolidinone carbonyl group was present regardless of the identity of the donor residue (which is Thr62 in GFP). Moreover, the imidazolidinone carbonyl group acts as an acceptor despite substantial diversity in the chemical structure of its chromophore (Fig. 1).

To discern whether one or both of the lone pairs is involved in electron donation to the electrophilic acceptor carbonyl group, we measured the C′=O···C angle (ξ) formed by the donor carbonyl group and the acceptor carbon. The electron lone pairs of the donor oxygen are oriented either along the carbonyl bond axis or orthogonal to that axis [Fig. 3(A,B)].[28–30] The lone pair oriented along the C=O axis is s-rich, whereas the other lone pair is primarily p-rich. Both of these lone pairs can donate their electron density to an empty π* orbital. If the angle between the donor carbonyl group and the acceptor carbon is ~180°, then the electron donation is primarily from the s-rich lone pair. In contrast, if the angle is ~90°, then the electron donation is mainly from the p-rich lone pair. Angles between these two extremes are indicative of electron donation from both of the lone pairs. We found that this angle has a mean and median of $\xi = 143°$ (Fig. 4, Supporting Information Table S1), indicating that substantial electron donation originates from both of the lone pairs. This origin contrasts with that for the n→π* interaction between the main-chain carbonyl groups of adjacent residues in an α-helix, which stems almost exclusively from the p-rich lone pair.[8,16]

Next, we resorted to Natural Bond Orbital (NBO)[31–33] analysis to estimate the strength of the n→π* electronic delocalization. The coordinates of the key functional groups were extracted from a high-resolution structure.[34] NBO analysis revealed that both lone pairs are poised to interact with the antibonding orbital (π*) of the acceptor carbonyl group [Fig. 3(C–E)]. In addition, second-order perturbation theory estimated the strength of the n→π* interaction in this structure to be $E_{n\rightarrow\pi^*} = 0.84$ kcal/mol, which is greater than that estimated for a typical n→π* interaction between adjacent residues in an α-helix ($E_{n\rightarrow\pi^*} \approx 0.5$ kcal/mol).[8]

Is there an experimental signature for the n→π* interaction with the chromophore of fluorescent proteins? Such an interaction should lower the frequency of the stretching vibration of the acceptor carbonyl group. In accord with this anticipation, infrared spectroscopy studies indicate that this

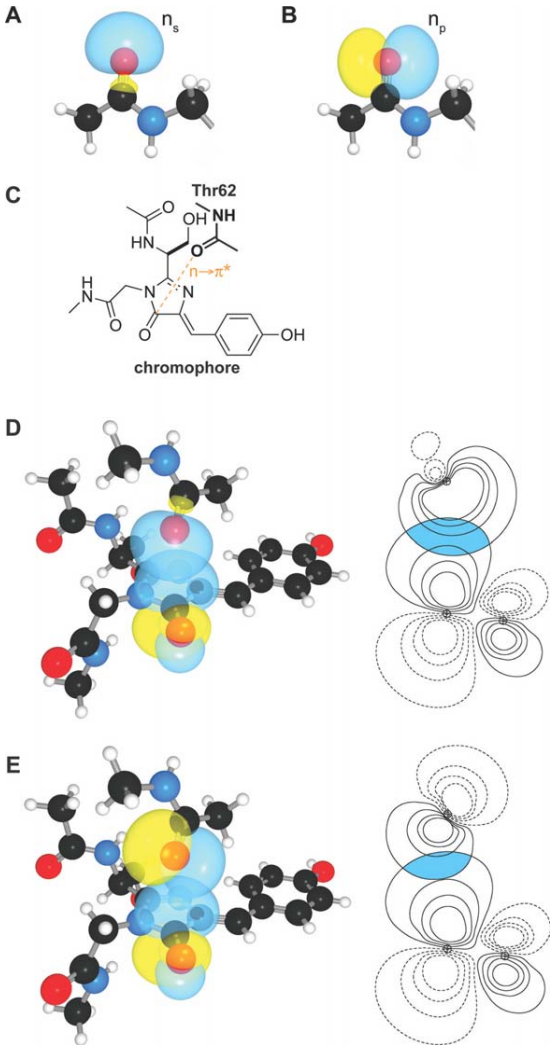

Figure 3. Lone pairs of an amide oxygen. (A) s-Rich lone pair, n_s. (B) p-Rich lone pair, n_p. (C) Two fragments of GFP derived from a structure determined at 0.9-Å resolution (PDB entry 2wur)[34] and used in computational analyses. (D) Overlap of n_s of the main-chain oxygen of Thr62. (E) Overlap of n_p of the main-chain oxygen of Thr62.

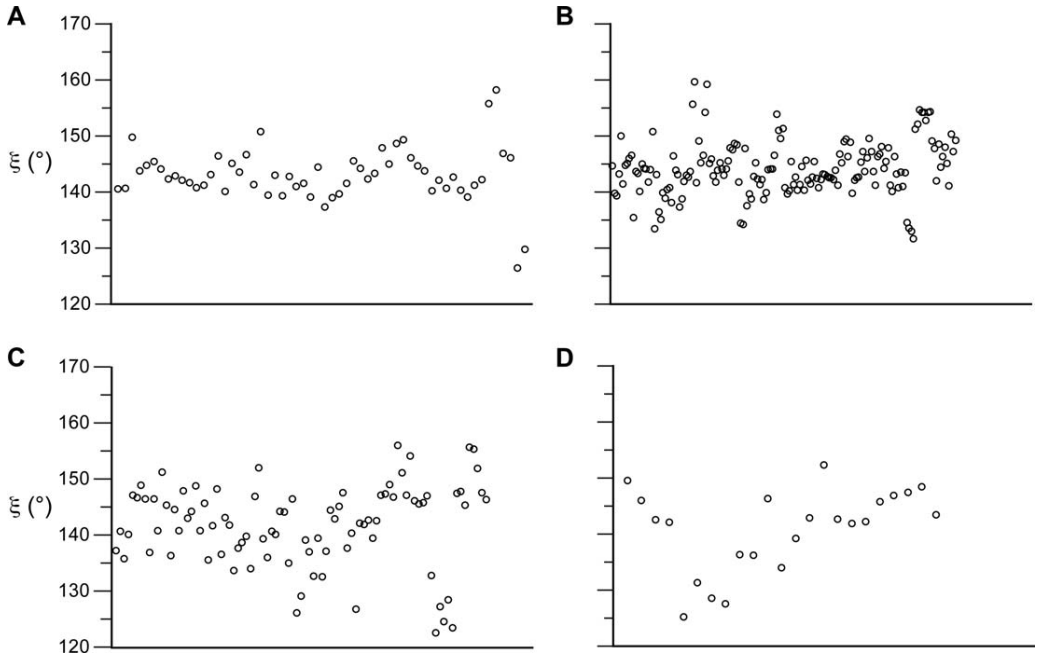

Figure 4. C′=O···C angle (ξ) for fluorescent protein structures determined at different resolutions: (A) <1.50 Å, (B) 1.50–1.99 Å, (C) 2.00–2.49 Å, and (D) ≥ 2.50 Å. The horizontal axis serves merely to distribute the data.

frequency is 15 cm^{-1} lower in GFP than in a model imidazolidinone in a nonprotein environment.[35]

The close proximity of a main-chain carbonyl group to the chromophore could play a key role in the manifestation of its fluorescence. The chromophores of model compounds have low quantum yields ($\Phi < 4 \times 10^{-4}$) compared to that of GFP ($\Phi = 0.8$).[36,37] The basis for such low quantum yields has been attributed to the facile rotation about the exocyclic carbon–carbon double bond (Fig. 1), which causes nonradiative decay of the excited state.[38–41] In these model compounds, the excited state can adopt a twisted conformation in which fluorescence quenching can occur through nonadiabatic crossing. We hypothesize that the oxygen of the donor carbonyl group restrains the imidazolidinone ring of the chromophore in the rigid state that is necessary for its unusual photophysical properties, deterring the attainment of the deleterious twisted conformation.

A high-resolution structure of GFP reveals another intimate interaction between residue 62 and the aromatic ring of the chromophore. In this structure,[42] a $C_{\gamma 2}$–H of Thr62 is situated over the phenolic ring so as to form a C—H···π interaction (Fig. 5). Like the n→π* interaction and other packing interactions, this C—H···π interaction could serve to maintain the rigidity of the chromophore. We speculate that analogous interactions operate in RNA mimics of GFP.[43]

The biogenesis of the GFP chromophore has been studied and debated intensely.[44,45] Although the mechanistic pathway for the maturation of the

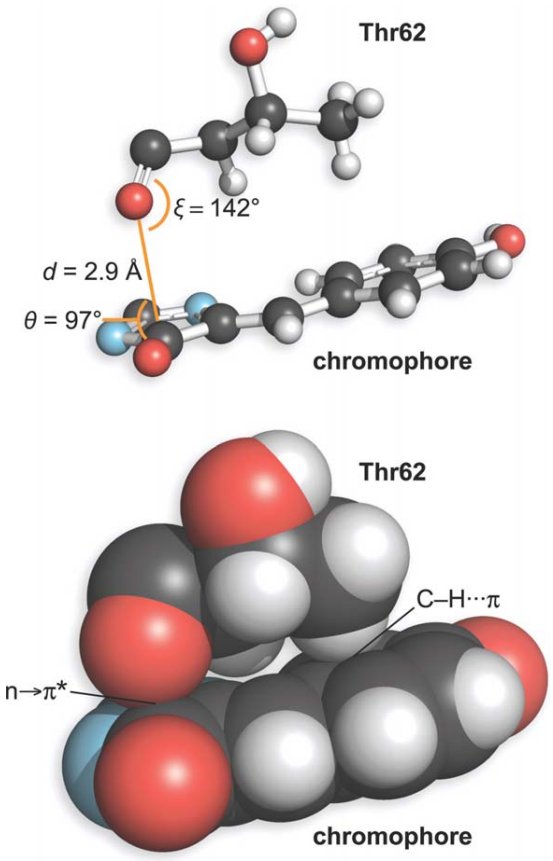

Figure 5. n→π* and C–H···π interactions in a structure of GFP determined at a resolution of a 0.9 Å (PDB entry 2wur).[34] (A) Ball-and-stick image. (B) Space-filling image.

chromophore is controversial, some evidence is attainable from the crystal structures of on-pathway intermediates in GFP variants with reduced rates of chromophore maturation (Fig. 6).[46–51] A number of crystal structures are known for the conformation of the precyclized state, and these structures are of two distinct types.

In one type of precyclized structure, an n→π* interaction exists between the main-chain carbonyl group of residue 62 and the chromophore [d = 3.0 Å, θ = 97°; Fig. 7(A)].[46] The main-chain residues are preorganized for the cyclization reaction,[46,45] and we postulate that the n→π* interaction assists in this preorganization. In addition, the n→π* interaction likely disturbs the amidic resonance of the main-chain amide that contains the nucleophilic nitrogen, making that nitrogen more like one in an amino group and endowing it with greater nucleophilicity.

In the other type of precyclized structure, the carbonyl group that is destined to be part of the imidazolidinone is involved in an n→π* interaction with that of an adjacent residue rather than with the carbonyl group of residue 62 [Fig. 7(B)].[48,49] This n→π* interaction disturbs the amidic resonance of the nucleophilic nitrogen [as in Fig. 7(A)] but also drains electron density from the donor carbonyl group, facilitating attack by that nitrogen. In other words, this n→π* interaction [Fig. 7(B)] could make the electrophilic and nucleophilic centers more reactive simultaneously. Two other on-pathway intermediates, namely the cyclized[47] and dehydrated structures,[51] have considerable n→π* interaction between the donor carbonyl group of residue 62 and the acceptor carbonyl group of the imidazolidinone ring [Figs. 7(C,D), Supporting Information Table S2].

The relative position of the donor carbonyl group with respect to the acceptor carbonyl group is unclear in the excited state.[52–54] Nonetheless, infrared spectroscopy suggests that the acceptor carbonyl stretching mode is preserved in the excited state, even though deprotonation of the phenolic oxygen has taken place.[35] If significant structural changes do not occur in the excited state, then the n→π* interaction could persist there. Accordingly, we encourage the inclusion of this interaction in computational and photophysical analyses of the chromophore, as well as in speculations about the molecular evolution of fluorescent proteins.

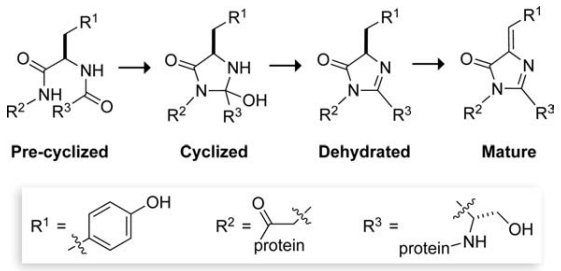

Figure 6. Putative mechanism of chromophore biogenesis in GFP.[46,47,51]

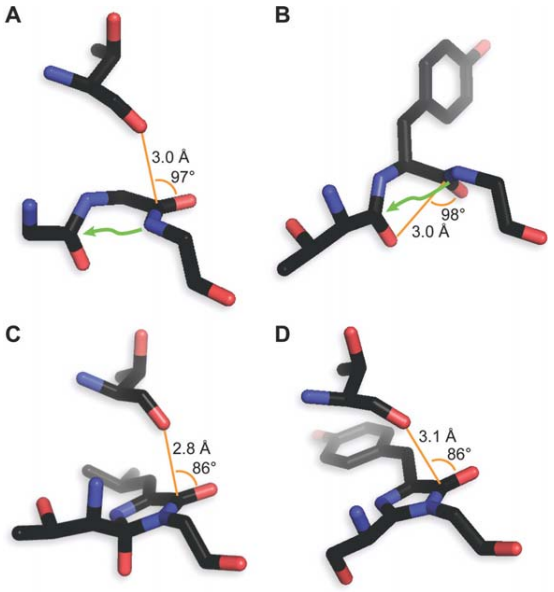

Figure 7. Intermediates in chromophore biogenesis in GFP. In panels A and B, the arrows indicate electron flow to form the cyclized intermediate. (A) Precyclized intermediate with n→π* interaction between residue 62 and the incipient chromophore (PDB entry 1qyo).[46] (B) Precyclized intermediate with n→π* interaction between adjacent residues in the incipient chromophore (PDB entry 1qy3).[46] (C) Cyclized intermediate (PDB entry 1s6z).[47] (D) Dehydrated intermediate (PDB entry 3lvc).[51]

Conclusion

Our database and computational analyses have revealed the existence of an n→π* interaction between a main-chain oxygen and the chromophore of fluorescent proteins. This interaction could contribute to the biogenesis and photophysical properties of the chromophore.

Methods

PDB analyses

We collected the atomic coordinates of fluorescent proteins that had been deposited in the Protein Data Bank (RCSB PDB) by March 1, 2011, and we grouped the datasets according to their resolution. The distance and angle between the donor and acceptor groups were measured with the program PyMOL (Schrödinger, Portland, OR). The measured distances and angles can be found in the Supporting Information.

Computational analyses

We used Natural Bond Orbital (NBO) analyses to estimate the strength of the n→π* electronic

delocalization.[31–33] The coordinates of the key functional groups were extracted from a structure determined to atomic resolution (PDB entry 2wur) and capping groups were installed at N- and C-termini to give the two GFP fragments shown in Figure 3(C). Natural bond orbital analyses on these fragments were performed at the B3LYP/6-311+G(2d,p) level of theory. The NBO method deconstructs a calculated wavefunction into a localized form, which corresponds to the lone pair and bond elements of the Lewis structure. (For additional details, see: http://www.chem.wisc.edu/~nbo5/web_nbo.htm.) The stabilization afforded by the n→π* electronic delocalization, $E_{n\rightarrow\pi^*}$, was estimated by using second-order perturbation theory as implemented in NBO 5.0. Orbital depictions were generated with the program NBOView 1.1.

Acknowledgments

We thank B.R. Caes for contributive discussions. K.J.K. was supported by a Barry M. Goldwater Scholarship and a Hilldale Undergraduate/Faculty Research Fellowship.

References

1. Tsien RY (1998) The green fluorescent protein. Annu Rev Biochem 67:509–544.
2. Zimmer M (2002) Green fluorescent protein (GFP): applications, structure, and related photophysical behavior. Chem Rev 102:759–781.
3. Chalfie M (2009) GFP: Lighting up life (Nobel lecture). Angew Chem Int Ed 48:5603–5611.
4. Shimomura O (2009) Discovery of green fluorescent protein (GFP) (Nobel lecture). Angew Chem Int Ed 48: 5590–5602.
5. Tsien RY (2009) Constructing and exploiting the fluorescent protein paintbox (Nobel lecture). Angew Chem Int Ed 48:5612–5626.
6. Chudakov DM, Matz MV, Lukyanov S, Lukyanov KA (2010) Fluorescent proteins and their applications in imaging living cells and tissues. Physiol Rev 90: 1103–1163.
7. Remington SJ (2011) Green fluorescent protein: a perspective. Protein Sci 20:1508–1519.
8. Bartlett GJ, Choudhary A, Raines RT, Woolfson DN (2010) n→π* interactions in proteins. Nat Chem Biol 6: 615–620.
9. DeRider ML, Wilkens SJ, Waddell MJ, Bretscher LE, Weinhold F, Raines RT, Markley JL (2002) Collagen stability: insights from NMR spectroscopic and hybrid density functional computational investigations of the effect of electronegative substituents on prolyl ring conformations. J Am Chem Soc 124:2497–2505.
10. Hinderaker MP, Raines RT (2003) An electronic effect on protein structure. Protein Sci 12:1188–1194.
11. Hodges JA, Raines RT (2006) Energetics of an n→π* interaction that impacts protein structure. Org Lett 8: 4695–4697.
12. Choudhary A, Gandla D, Krow GR, Raines RT (2009) Nature of amide carbonyl–carbonyl interactions in proteins. J Am Chem Soc 131:7244–7246.
13. Choudhary A, Fry CG, Raines RT (2010) Modulation of an n→π* interaction with α-fluoro groups. ARKIVOC 8: 251–262.
14. Jakobsche CE, Choudhary A, Miller SJ, Raines RT (2010) n→π* interaction and n)(π Pauli repulsion are antagonistic for protein stability. J Am Chem Soc 132: 6651–6653.
15. Shoulders MD, Kotch FW, Choudhary A, Guzei IA, Raines RT (2010) The aberrance of the 4*S* diastereomer of 4-hydroxyproline. J Am Chem Soc 132:10857–10865.
16. Choudhary A, Raines RT (2011) Signature of n→π* interactions in α-helices. Protein Sci 20:1077–1081.
17. Bürgi HB, Dunitz JD, Shefter E (1973) Geometrical reaction coordinates. II. Nucleophilic addition to a carbonyl group. J Am Chem Soc 95:5065–5067.
18. Bürgi HB, Dunitz JD, Lehn JM, Wipff G (1974) Stereochemistry of reaction paths at carbonyl centres. Tetrahedron 30:1563–1572.
19. Bürgi HB, Lehn JM, Wipff G (1974) An *ab initio* study of nucleophilic addition to a carbonyl group. J Am Chem Soc 96:1965–1966.
20. Blanco S, López JC, Mata S, Alonso JL (2010) Conformations of γ-aminobutyric acid (GABA): the role of the n→π* interaction. Angew Chem Int Ed 49: 9187–9192.
21. Choudhary A, Kamer KJ, Raines RT (2011) An n→π* interaction in aspirin: implications for structure and reactivity. J Org Chem 76:7933–7937.
22. Gorske BC, Bastian BL, Geske GD, Blackwell HE (2007) Local and tunable n→π* interactions regulate amide isomerism in the peptoid backbone. J Am Chem Soc 129:8928–8929.
23. Gorske BC, Stringer JR, Bastian BL, Fowler SA, Blackwell HE (2009) New strategies for the design of folded peptoids revealed by a survey of noncovalent interactions in model systems. J Am Chem Soc 131: 16555–16567.
24. Choudhary A, Kamer KJ, Powner MW, Sutherland JD, Raines RT (2010) A stereoelectronic effect in prebiotic nucleotide synthesis. ACS Chem Biol 5:655–657.
25. Ormö M, Cubitt AB, Kallio K, Gross LA, Tsien RY, Remington SJ (1996) Crystal structure of the *Aequorea victoria* green fluorescent protein. Science 273: 1392–1395.
26. Yang F, Moss LG, Phillips GN, Jr. (1996) The molecular structure of green fluorescent protein. Nat Biotechnol 14:1246–1251.
27. Berman H, Henrick K, Nakamura H, Markley JL (2007) The worldwide protein data bank (wwPDB): ensuring a single, uniform archive of PDB data. Nucleic Acids Res 35:D301–D303.
28. Gray HB. 1965. Electrons and chemical bonding. New York: W. A. Benjamin.
29. Raber DJ, Raber NK, Chandrasekhar J, Scheleyer PvR (1984) Geometries and energies of complexes between formaldehyde and first- and second-row cations. A theoretical study. Inorg Chem 23:4076–4080.
30. Laing M (1987) No rabbit ears on water. J Chem Educ 64:124–128.
31. Weinhold F. Natural bond orbital methods. In: Schleyer PvR, Allinger NL, Clark T, Gasteiger J, Kollman PA, Shaefer HF, III, Schreiner PR, editors. (1998) Encyclopedia of computational chemistry. Chichester, UK: Wiley, pp 1792–1811.
32. Glendening ED, Badenhoop JK, Reed AE, Carpenter JE, Bohmann JA, Morales CM, Weinhold F (2001) NBO 5.0. NBO 50.
33. Weinhold F, Landis CR (2005) Valency and bonding: a natural bond orbital donor–acceptor perspective. Cambridge, UK: Cambridge University Press.
34. Shinobu A, Palm GJ, Schierbeek AJ, Agmon N (2010) Visualizing proton antenna in a high-resolution green

fluorescent protein structure. J Am Chem Soc 132: 11093–11102.
35. Stoner-Ma D, Melief EH, Nappa J, Ronayne KL, Tonge PJ, Meech SR (2006) Proton relay reaction in green fluorescent protein (GFP): polarization-resolved ultrafast vibrational spectroscopy of isotopically edited GFP. J Phys Chem B 110:22009–22018.
36. Niwa H, Inouye S, Hirano T, Matsuno T, Kojima S, Kubota M, Ohashi M, Tsuji FI (1996) Chemical nature of the light emitter of the *Aequorea* green fluorescent protein. Proc Natl Acad Sci USA 93:13617–13622.
37. Wu L, Burgess K (2008) Synthesis of highly fluorescent GFP-chromophore analogues. J Am Chem Soc 130: 4089–4096.
38. Voityuk AA, Michel-Beyerle M-B, Rösch N (1998) Quantum chemical modeling of structure and absorption spectra of the chromophore in green fluorescent proteins. Chem Phys 231:13–25.
39. Weber W, Helms V, McCammon JA, Langhoff PW (1999) Shedding light on the dark and weakly fluorescent states of green fluorescent proteins. Proc Natl Acad Sci USA 96:6177–6182.
40. Litvinenko KL, Webber NM, Meech SR (2001) An ultrafast polarisation spectroscopy study of internal conversion and orientational relaxation of the chromophore of the green fluorescent protein. Chem Phys Lett 346:47–53.
41. Webber NM, Litvinenko KL, Meech SR (2001) Radiationless relaxation in a synthetic analogue of the green fluorescent protein chromophore. J Phys Chem B 105: 8036–8039.
42. Brandl M, Weiss MS, Jabs A, Suhnel J, Hilgenfeld R (2001) C–H···π interactions in proteins. J Mol Biol 307: 357–377.
43. Paige JS, Wu KY, Jaffrey SR (2011) RNA mimics of green fluorescent protein. Science 333:642–646.
44. Remington SJ (2006) Fluorescent proteins: maturation, photochemistry and photophysics. Curr Opin Struct Biol 16:714–721.
45. Wachter RM (2007) Chromogenic cross-link formation in green fluorescent protein. Acc Chem Res 40: 120–127.
46. Barondeau DP, Putnam CD, Kassmann CJ, Taines JA, Getzoff ED (2003) Mechanism and energetics of green fluorescent protein chromophore synthesis revealed by trapped intermediate structures. Proc Natl Acad Sci USA 100:12111–12116.
47. Rosenow MA, Huffman HA, Phail ME, Wachter RM (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.
48. Barondeau DP, Kassmann CJ, Tainer JA, Getzoff ED (2005) Understanding GFP chromophore biosynthesis: controlling backbone cyclization and modifying post-translational chemistry. Biochemistry 44:1960–1970.
49. Wood TI, Barondeau DP, Hitomi C, Kassmann CJ, Tainer JA, Getzoff ED (2005) Defining the role of arginine 96 in green fluorescent protein fluorophore biosynthesis. Biochemistry 44:16211–16220.
50. Barondeau DP, Tainer JA, Getzoff ED (2006) Structural evidence for an enolate intermediate in GFP fluorophore biosynthesis. J Am Chem Soc 128:3166–3168.
51. Pletneva NV, Pletnev VZ, Lukyanov KA, Gurskaya NG, Goryacheva EA, Martynov VI, Wlodawer A, Dauter Z, Pletnev S (2010) Structural evidence for a dehydrated intermediate in green fluorescent protein chromophore biosynthesis. J Biol Chem 285:15978–15984.
52. Olsen S, Smith SC (2007) Radiationless decay of red fluorescent protein chromophore models via twisted intramolecular charge-transfer states. J Am Chem Soc 129:2054–2065.
53. Olsen S, Smith SC (2008) Bond selection in the photoisomerization reaction of anionic green fluorescent protein and kindling fluorescent protein chromophore models. J Am Chem Soc 130:8677–8689.
54. Meech SR (2009) Excited state reactions in fluorescent proteins. Chem Soc Rev 38:2922–2934.