Monatsh Chem (2011) 142:709–715
DOI 10.1007/s00706-010-0430-1

REVIEW

Absorption tuning of the green fluorescent protein chromophore: synthesis and studies of model compounds

Mogens Brøndsted Nielsen · Lars H. Andersen · Tomás Rocha-Rinza

Received: 25 October 2010 / Accepted: 18 November 2010 / Published online: 16 December 2010

Abstract The green fluorescent protein (GFP) chromophore is a heterocyclic compound containing a *p*-hydroxybenzylidine attached to an imidazol-5(4*H*)-one ring. This review covers the synthesis of a variety of model systems for elucidating the intrinsic optical properties of the chromophore in the gas phase and its response in particular to hydrogen bond interactions. The overall goal is to understand how the protein binding pocket influences the absorption behavior, and the current status of our ongoing efforts is presented.

Keywords Heterocycles · Hydrogen bonds · Supramolecular chemistry · Protein interactions · Absorption tuning

Introduction

Heterocyclic chromophores are widely distributed in Nature. Examples include chlorophylls involved in light harvesting during photosynthesis; phytochromes as regulators of photomorphogenesis in plants; anthocyanins responsible for the red, purple, and blue colors in many fruits, vegetables, cereal grains, and flowers; flavins which mediate light-induced responses in plants, including phototropism, movements of chloroplasts and leaves, and onset of flowering (for a review on plant photoreceptors, see [1]). Together with another heterocyclic chromophore, methenyltetrahydrofolylpolyglutamate (MTHF), flavin is responsible for the function of some DNA photolyases (for DNA repair). In water MTHF absorbs at 358 nm, but inside the protein its absorption is different. Thus, the protein environment in the DNA photolyase of *E. coli* induces a redshift in the MTHF absorption maximum to 383 nm, at which wavelength the flux of solar radiation is higher and in this way MTHF becomes a more efficient light antenna [2]. The protein environment hence plays an important role for tuning the optical properties of the chromophore. Albeit not a heterocyclic chromophore, the pronounced absorption tuning of the retinal chromophore also deserves mention. By embedding the same retinal chromophore in different protein environments (linked to the protein as a protonated Schiff base), the entire visual absorption spectrum can be covered, which allows us to distinguish colors in the process of vision [3].

The green fluorescent protein (GFP) is another light-sensitive protein that was first discovered in 1962 in the jellyfish *Aequorea victoria* [4]. It converts blue light into green light and has found wide applications as a marker protein for molecular and cell biology owing to these fluorescent properties [5–11]. The covalently linked chromophore inside the protein is a *p*-hydroxybenzylidine-imidazolinone, formed autocatalytically from residues serine 65, tyrosine 66, and glycine 67. It is embedded in a network of hydrogen bonds inside the GFP binding pocket [12] (Fig. 1). These hydrogen bond interactions may play a role in the absorption characteristics of the chromophore as indicated by several mutation studies [13–15]. The absorption spectrum of GFP shows two maxima, one at

M. B. Nielsen (✉)
Department of Chemistry, University of Copenhagen, 2100 Copenhagen Ø, Denmark
e-mail: mbn@kiku.dk

L. H. Andersen
Department of Physics and Astronomy, University of Aarhus, 8000 Århus C, Denmark

T. Rocha-Rinza
Instituto de Química, UNAM, Circuito Exterior Ciudad Universitaria, Delegación Coyoacán, C.P. 04510 Mexico City, Mexico

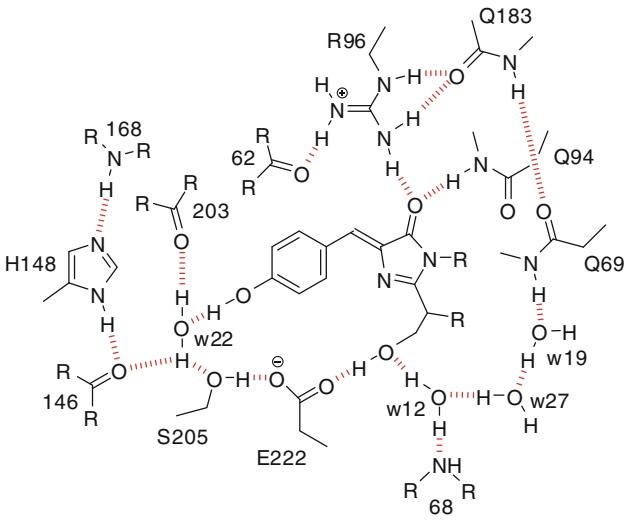


Fig. 1 Binding pocket for the GFP chromophore. Adapted from Ref. [12]

479 nm and one at 397 nm. The first absorption is ascribed to the deprotonated chromophore (phenolate anion), whereas the second is ascribed to the neutral phenol.

Of particular interest is the challenge to discover how the absorption properties of the neutral and anionic GFP chromophores are tuned by protein interactions. Our approach is based on gas-phase spectroscopy of suitable compounds, each of which models one specific interaction. We aim at studying the optical properties in the absence of an interfering medium like a solvent or a protein. In this review, the synthesis of model compounds of the anionic and neutral GFP chromophore and their intrinsic optical properties will be presented. First, we will look at absorption tuning of the phenolate, and next, at the neutral phenol with a special emphasis on the hydrogen bonding to

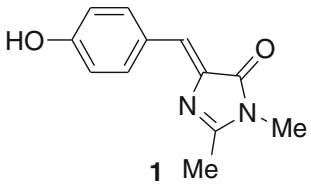


Fig. 3 Neutral GFP chromophore

the carbonyl oxygen. A variety of model systems designed for elucidating the influence of hydrogen bonding and/or the approach of a positive charge to the carbonyl moiety of the chromophore are depicted in Fig. 2.

Gas-phase spectroscopy of charged species

The gas-phase absorption spectrum is measured by action spectroscopy employing an electrostatic ion storage ring in combination with advanced laser setup as described in detail elsewhere [16–19]. In short, ions are injected into the storage ring after electrospray ionization, acceleration, and mass-to-charge selection by a bending magnet. The electrospray ionization technique allows the investigation of charged, nonvolatile species such as the GFP phenolate. The ion bunch is excited by a pulsed laser, which leads to fragmentation. The number of neutral fragments is measured and correlated to the absorption cross section at a specific wavelength. In other words, the more neutral fragments, the more photons of a given wavelength had been absorbed. The technique does not allow studies on neutral molecules like the true neutral GFP chromophore **1** (Fig. 3) as it would become charged under the conditions of electrospray ionization. Instead, a "spectator charge" needs to be attached, for example, an ammonium group (cf. Fig. 2).

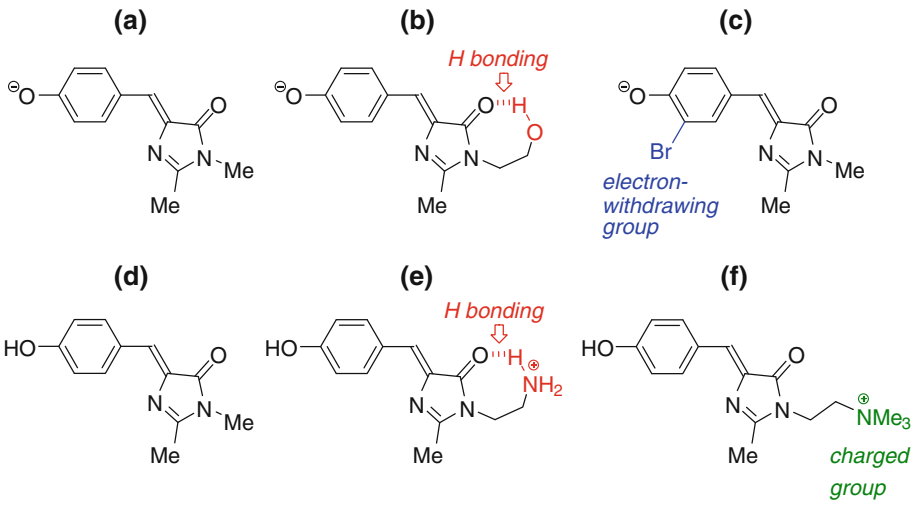
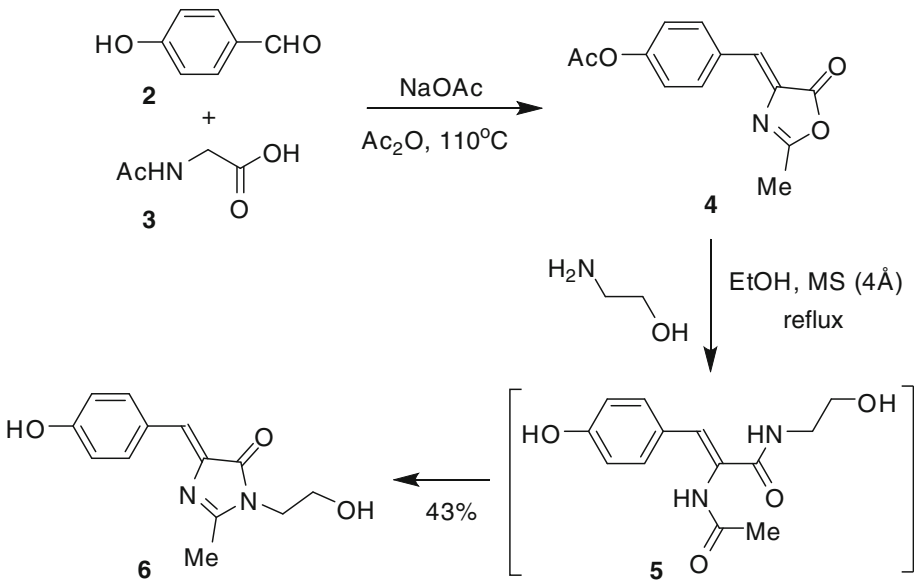
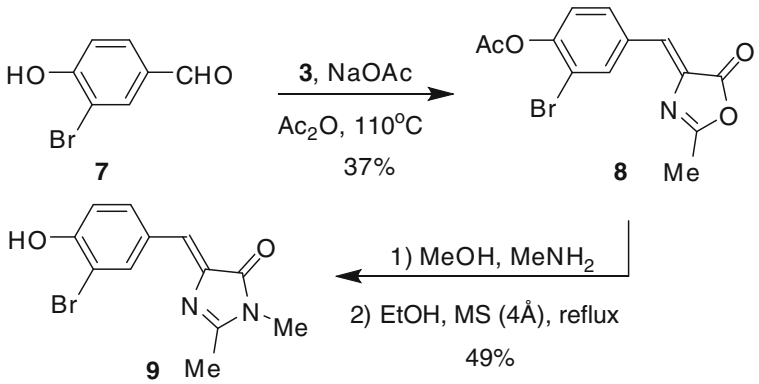


Fig. 2 Model systems for elucidating absorption tuning in the anionic (**a**–**c**) and neutral (**d**–**f**) GFP chromophore

Scheme 1



Scheme 2



Results and discussion

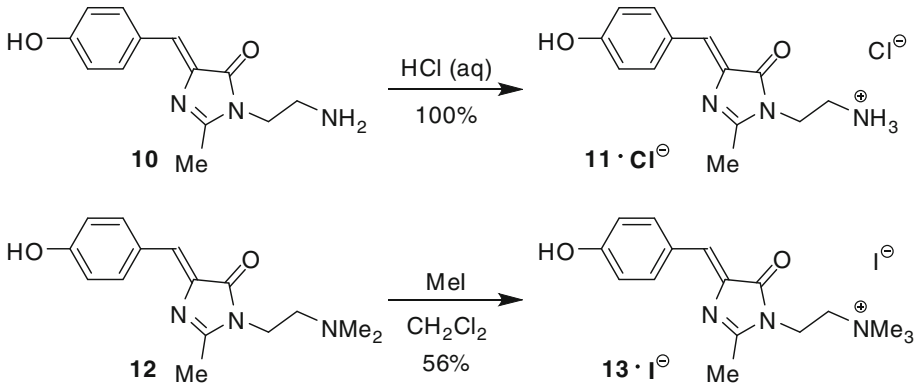
Synthesis of GFP chromophores

The synthesis of a model system incorporating a hydroxyethyl arm for hydrogen bonding to the carbonyl oxygen is shown in Scheme 1 [20]. First, *p*-hydroxybenzaldehyde (**2**) and *N*-acetylglycine (**3**) were condensed to form the oxazolone **4**. Treatment of this compound with 2-aminoethanol in refluxing ethanol in the presence of molecular sieves (MS) resulted in nucleophilic attack onto the carbonyl followed by ring opening to provide the intermediate **5**, for which deacetylation had also occurred. Under the reaction conditions this compound subsequently underwent ring closure and dehydration to form the imidazolone target compound **6**.

In a similar manner, 3-bromo-4-hydroxybenzaldehyde (**7**) was converted to the oxazolone **8** which was subsequently treated with methylamine to ultimately furnish the product **9** incorporating an electron-withdrawing bromo substituent at the phenol ring (Scheme 2) [21].

GFP chromophores with amino functionalities were prepared under similar conditions by treating the oxazolone **4** with suitable diamines [20]. Thus, we prepared the amine **10** that was subsequently precipitated as its hydrochloride salt **11**·Cl^- (Scheme 3). The dimethylamine **12** was also prepared and, rather than being treated with aqueous HCl, it was methylated with methyl iodide to provide the salt **13**·I^- (Scheme 3). The properties of these two compounds are important for comparison as only **11** is able to form a hydrogen bond with the carbonyl

Scheme 3



oxygen, while the influence of the positive charge may be inferred from **13**.

Absorption properties

The phenolates exhibit absorption maxima around 426–430 nm in methanol solution [21]. Thus, the polar solvent essentially averages out any potential influence exerted by the different functionalizations. The phenolates were subjected to gas-phase action spectroscopy after electrospray ionization of the corresponding phenols from a methanol solution [21, 22]. Figure 4 shows the gas-phase absorption spectra of the two isolated phenolates of **6** and **9** and, for comparison, that of the "parent phenolate" of **1** [21]. Somewhat surprisingly, the three chromophores exhibit similar absorption maxima at 482 ± 2 nm. Thus, hydrogen bonding to the carbonyl oxygen from a neutral hydrogen donor does not influence the absorption. By using density functional theory (DFT) calculations we confirmed that the hydrogen-bonded conformation is indeed the most favorable, with a distance between the carbonyl oxygen and the hydroxy proton of 1.732 Å. The presence of a bromo substituent in the phenolate ring does not exert an effect either. The absorptions of the chromophores are, however, redshifted relative to the absorptions in solution. Thus, the absorption is clearly sensitive to solvent interactions (for a detailed study of the solvatochromism of the GFP chromophore, see [23]).

Interestingly, the intrinsic optical absorption of the phenolates is similar to that observed for the protein (479 nm). This observation may be interpreted in two ways: (1) interactions between the chromophore and the protein binding pocket cancel out, or (2) each individual interaction does not exert any influence on the chromophore absorption. As the chromophore is unaffected by a neutral hydrogen bond donor (**6** vs. **1**), it is reasonable to put forward that the neutral glutamine residue Q94 in GFP is unlikely to influence the absorption of the chromophore. We cannot exclude, however, absorption tuning by the positively charged arginine residue R96, as suggested from

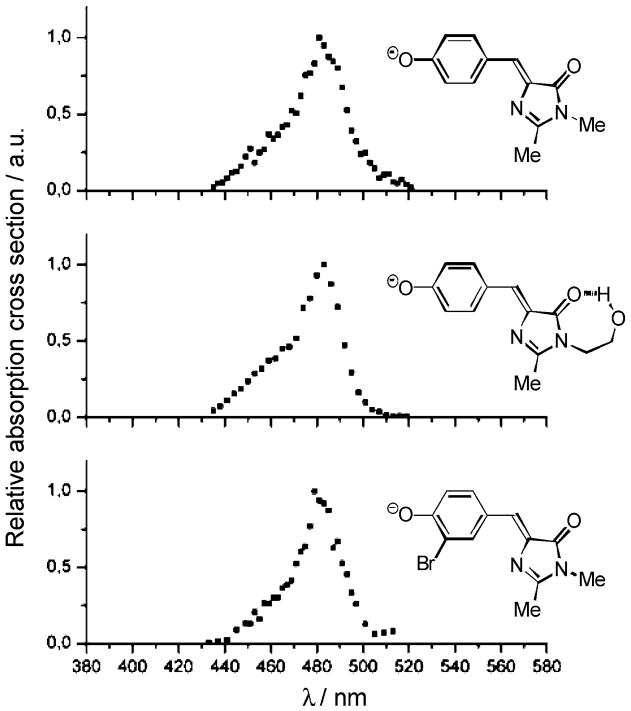


Fig. 4 Gas-phase absorption spectra of GFP phenolate chromophores derived from **1** (*top*), **6** (*middle*), and **9** (*bottom*). Figure taken from Ref. [21]. *Reproduced by permission of The Royal Society of Chemistry* (article link: doi:10.1039/b920378h)

mutation studies [13–15]. In addition, the effect of hydrogen bonding to the phenolate oxygen still needs to be unraveled. The effect of a positively charged residue on the photoresponse properties of the chromophores under consideration becomes evident when proceeding now to the gas-phase absorption properties of the neutral chromophores incorporating a charged "spectator ion".

The gas-phase studies on the GFP phenol chromophore include not only compounds **11** and **13**, but also compounds **14**–**17** (Fig. 5), which were prepared in parallel by Tolbert, Solntsev, and co-workers [24]. Altogether this series of compounds allows us to elucidate both the effect of hydrogen bonding between the carbonyl oxygen and a

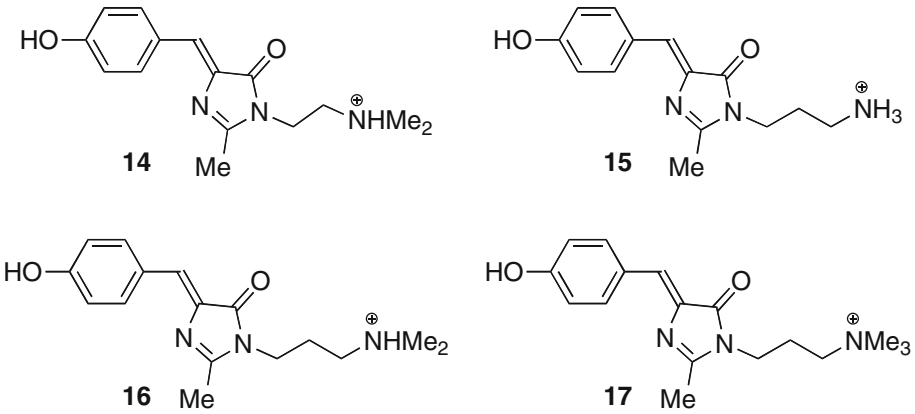


Fig. 5 Model systems prepared by Tolbert, Kolntsev, and co-workers [24]

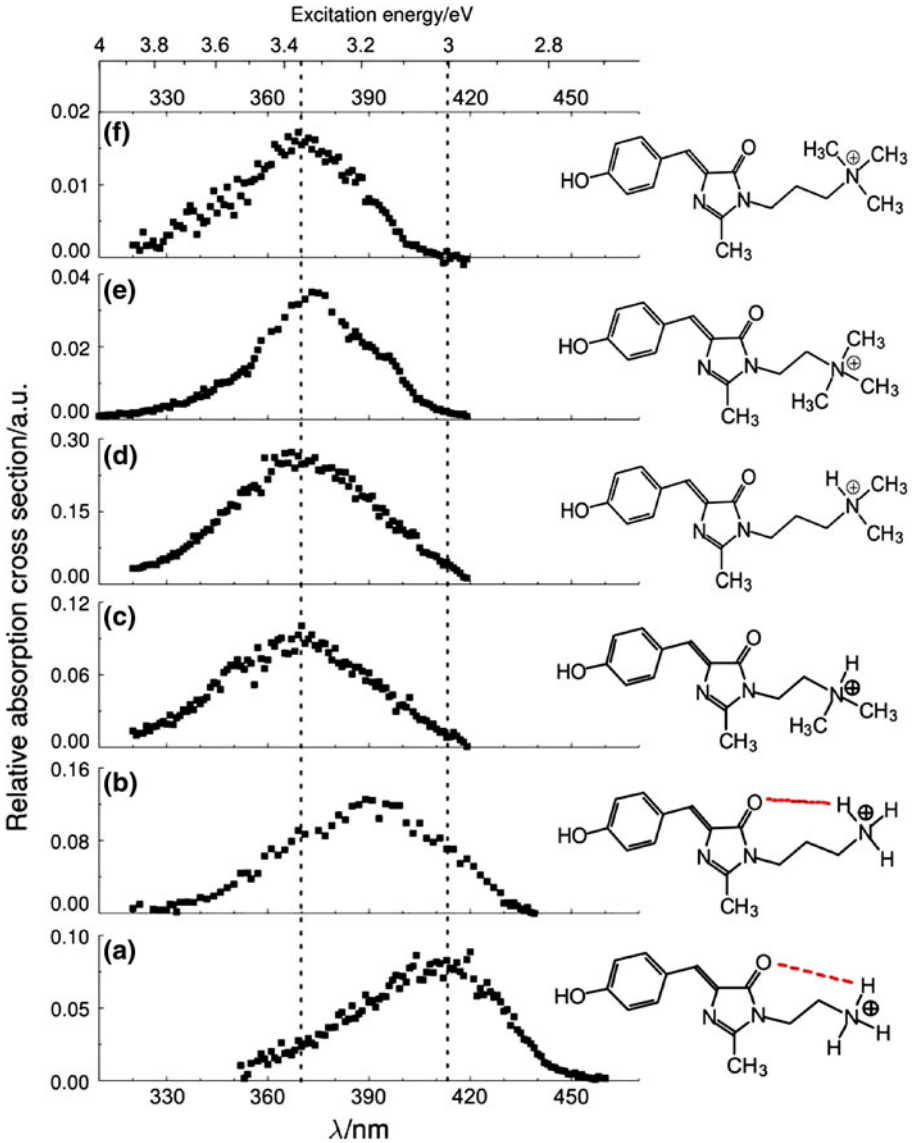
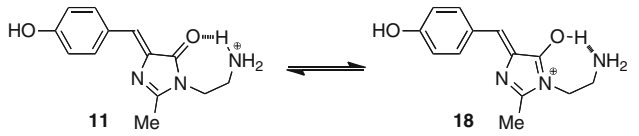


Fig. 6 Gas-phase absorption spectra of neutral GFP chromophore models incorporating a positively charged side group. Spectra from *top* to *bottom*: **17** (**f**), **13** (**e**), **16** (**d**), **14** (**c**), **15** (**b**), **11** (**a**) (in fact **18**, see text). Figure taken from Ref. [24]. *Reproduced by permission of the PCCP Owner Societies* (article link: doi: 10.1039/b914276b)

positively charged ammonium group and the influence of the charge itself at different distances. Gas-phase absorption spectra of the entire series are shown in Fig. 6 [24, 25].

First, we note that four of the chromophores (**13**, **14**, **16**, **17**) have their absorption maximum at about 370 nm, whereas the two others have their absorption maximum at



Scheme 4

394 nm (**15**) and 410 nm (**11**). The similar absorptions of cations **13** and **17**, both unable to form a hydrogen bond, show that increasing the distance to the positive charge from two to three methylene units in the spacer causes no change in the absorption maximum. The absorption redshift (ca. 25 nm) experienced by **15** clearly shows the effect of forming a hydrogen bond between the carbonyl oxygen and a charged hydrogen bond donor. Interestingly, however, the absorption is further redshifted for **11** in which the spacer is shortened to an ethylene unit. Quantum chemical calculations reveal that in fact compound **11** is expected to undergo a tautomerization reaction in the gas phase to the more stable tautomer **18** [by 0.8 kJ mol^{-1} according to the MP2 approximation with the basis set augmented with diffuse functions on the oxygen and nitrogen atoms, (aug)-cc-pVTZ] (Scheme 4) [24]. This tautomerization was also supported by excitation energy calculations at different levels (Table 1). Excellent agreement was observed between experimental and calculated excitation energies for the series of chromophores. Unfortunately, however, calculations on the true neutral chromophore **1** deviate significantly from each other. No experimental value can be used for comparison in this case (the spectator ion is required in our experimental technique). By aug-MCQDPT2/aug-cc-pVDZ a value of 399 nm is obtained, whereas a value of 349 nm is obtained by RI-CC2/aug-cc-pVDZ. The first value implies that the positive spectator charge in **13**, **14**, **16**, and **17** induces a blueshift in the absorption, whereas the second calculated value implies a redshift. Let us now compare these values to the absorption of the protein. The protein absorbs at 397 nm, and the binding pocket hence causes a redshift of almost 50 nm under the assumption that the isolated, neutral chromophore has an absorption maximum nearby 349 nm (RI-CC2). Such a large shift would be surprising in view of the rather small shifts that have been observed in GFP mutation studies [13–15]. If instead the value of 399 nm obtained by the aug-MCQDPT2 method is closer to the correct one, then the net spectral tuning on the neutral chromophore inside the protein is essentially zero as it is for the deprotonated (phenolate) chromophore. The latter scenario seems most likely, but clearly more calculations and studies (for example, on chromophores with the spectator charge more remotely located) are needed to clarify this situation.

Conclusions

A series of model compounds for the GFP chromophore was prepared via oxazolone intermediates. The method allowed incorporation of a group capable of forming a hydrogen bond to the carbonyl oxygen. Gas-phase studies revealed that the optical absorption of the phenolate chromophore is the same whether it is in vacuum or inside the protein binding pocket. Moreover, hydrogen bonding to the carbonyl oxygen by a neutral hydrogen donor does not exert any influence, which means that hydrogen bonding to the glutamine residue Q94 in GFP is not important for spectral tuning. For studying the neutral chromophore in the gas phase, a "spectator ion" is needed to allow for electrospray ionization. The experimental studies reveal that hydrogen bonding to the carbonyl oxygen by a charged ammonium group results in a redshifted absorption of ca. 25 nm relative to a chromophore incorporating a charged group unable to form a hydrogen bond. It may be difficult to compare this result directly with the GFP binding pocket, but it is interesting in view of the fact that a hydrogen bond does exist between the carbonyl oxygen and a protonated arginine (R96) in GFP (Fig. 1). Yet, calculations point in two opposite directions in regard to the sign of the absorption shift induced by hydrogen bonding relative to a true neutral chromophore. Elucidation of the absorption of the true neutral chromophore awaits both further experimental and calculational studies, and is important for stating the net effect exerted by the GFP environment on the chromophore. One of the model compounds underwent a tautomerization reaction in the gas phase by which the hydrogen-bonded proton was transferred from the ammonium group onto the carbonyl

Table 1 Experimental and calculated S_0–S_1 excitation energies (nm) for neutral GFP model chromophores in gas phase [24, 25]

	Influence of H-bond			Influence of positive charge				True neutral
	11	**18**	**15**	**14**	**16**	**13**	**17**	**1**
Experiment		410 ± 5	395 ± 5	370 ± 5	370 ± 5	370 ± 5	370 ± 5	
Aug-MCQDPT2/aug-cc-pVDZ	388	410	387	367		363		399
RI-CC2/aug-cc-pVDZ	396	406	394	383		365		349

oxygen. This tautomerization was supported by calculations and caused a redshifted absorption maximum.

Acknowledgments We are grateful to all co-workers whose names appear in the publications cited below. The Danish Research Agency (FNU) is gratefully acknowledged for financial support.

References

1. Möglich A, Yang X, Ayers RA, Moffat K (2010) Annu Rev Plant Biol 61:21
2. Henry AA, Jimenez R, Hanway D, Romesberg FE (2004) Chem Bio Chem 5:1088
3. Nielsen MB (2009) Chem Soc Rev 38:913
4. Shimomura FH, Johnson FH, Saiga Y (1962) J Cell Comp Physiol 59:223
5. Griffin BA, Adams SR, Tsien RY (1998) Science 281:269
6. Tsien RY (1998) Annu Rev Biochem 67:509
7. Zimmer M (2002) Chem Rev 102:759
8. Pakhomov AA, Martynov VI (2008) Chem Biol 15:755
9. Sanders JM, Jackson SE (2009) Chem Soc Rev 38:2821
10. Seward HE, Bagshaw CR (2009) Chem Soc Rev 38:2842
11. Craggs TD (2009) Chem Soc Rev 38:2865
12. Brejc K, Sixma TK, Kitts PA, Kain SR, Tsien RY, Ormö M, Remington SJ (1997) Proc Natl Acad Sci U S A 94:2306
13. Sniegowski JA, Phail ME, Wachter RM (2005) Biochem Biophys Res Commun 332:657
14. Jung G, Wiehler J, Zumbusch A (2005) Biophys J 88:1932
15. Wood TI, Barondeau DP, Hitomi C, Kassmann CJ, Tainer JA, Getzoff ED (2005) Biochemistry 44:16211
16. Møller SP (1997) Nucl Instrum Methods Phys Res A 394:281
17. Andersen JU, Hvelplund P, Nielsen SB, Tomita S, Wahlgreen H, Møller SP, Pedersen UV, Forster JS, Jørgensen TJD (2002) Rev Sci Instrum 73:1284
18. Andersen LH, Heber O, Zajfman D (2004) J Phys B At Mol Opt Phys 37:R59
19. Nielsen SB, Andersen JU, Hvelplund P, Liu B, Tomita S (2004) J Phys B At Mol Opt Phys 37:R25
20. Petersen MÅ, Riber P, Andersen LH, Nielsen MB (2007) Synthesis 3635
21. Lincke K, Sølling T, Andersen LH, Klærke B, Rahbek DB, Rajput J, Oehlenschlæger CB, Petersen MÅ, Nielsen MB (2010) Chem Commun 46:734
22. Nielsen SB, Lapierre, Andersen JU, Pedersen UV, Tomita S, Andersen LH (2001) Phys Rev Lett 87:2281021
23. Dong J, Solntsev KM, Tolbert LM (2006) J Am Chem Soc 128:12038
24. Rajput J, Rahbek DB, Andersen LH, Rocha-Rinza T, Christiansen O, Bravaya KB, Erokhin AV, Bochenkova AV, Solntsev KM, Dong J, Kowalik J, Tolbert LM, Petersen MÅ, Nielsen MB (2009) Phys Chem Chem Phys 11:9996
25. Lammich L, Petersen MÅ, Nielsen MB, Andersen LH (2007) Biophys J 92:201