Computational Prediction of Absorbance Maxima for a Structurally Diverse Series of Engineered Green Fluorescent Protein Chromophores

Qadir K. Timerghazin, Haley J. Carlson, Chen Liang, Robert E. Campbell,* and Alex Brown*

University of Alberta, Department of Chemistry, Edmonton, Alberta T6G 2G2, Canada

Received: October 10, 2007; In Final Form: November 15, 2007

By virtue of its self-sufficiency to form a visible wavelength chromophore within the confines of its tertiary structure, the *Aequorea victoria* green fluorescent protein (GFP) is single-handedly responsible for the ever-growing popularity of fluorescence imaging of recombinant fusion proteins in biological research. Engineered variants of GFP with altered excitation or emission wavelength maxima have helped to expand the range of applications of GFP. The engineering of the GFP variants is usually done empirically by genetic modifications of the chromophore structure and/or its environment in order to find variants with new photophysical properties. The process of identifying improved variants could be greatly facilitated if augmented or guided by computational studies of the chromophore ground and excited-state properties and dynamics. In pursuit of this goal, we now report a thorough investigation of computational methods for prediction of the absorbance maxima for an experimentally validated series of engineered GFP chromophore analogues. The experimental dataset is composed of absorption maxima for 10 chemically distinct GFP chromophore analogues, including a previously unreported Y66D variant, measured under identical denaturing conditions. For each chromophore analogue, excitation energies and oscillator strengths were calculated using configuration interaction with single excitations (CIS), CIS with perturbative correction for double substitutions [CIS(D)], and time-dependent density functional theory (TD DFT) using several density functionals with solvent effects included using a polarizable continuum model. Comparison of the experimental and computational results show generally poor quantitative agreement with all methods attempted. However, good linear correlations between the calculated and experimental excitation energies ($R^2 > 0.9$) could be obtained. Oscillator strengths obtained with TD DFT using pure density functionals also correlate well with the experimental values. Interestingly, most of the computational methods used in this work fail in the case of nonaromatic Y66S and Y66L protein chromophores, which may be related to a significant contribution of double excitations to their excited-state wavefunctions. These results provide an important benchmark of the reliability of the computational methods as applied to GFP chromophore analogues and lays a foundation for the computational design of GFP variants with improved properties for use in biological imaging.

1. Introduction

Fluorescent proteins (FPs), defined as homologues and variants of the *Aequorea* jellyfish green FP (GFP), are popular and ubiquitous tools in molecular and cellular biology.[1] The utility of this expanding class of proteins stems from their shared ability to autocatalytically generate a visible wavelength chromophore within the confines of their three-dimensional structure (Figure 1). This unique characteristic has led to the extensive use of FPs as nonobtrusive reporter molecules for real-time fluorescence imaging and tracking of protein chimeras in transfected cells and transgenic organisms. Through the efforts of protein engineers, the FP 'toolbox' is now filled with a palette of variants with absorption and emission maxima that span most of the visible spectrum. By selecting appropriate variants from the toolbox it is now possible to track multiple protein chimeras or detect protein−protein interaction by Förster (or fluorescence) resonance energy transfer (FRET) in live cells.

The chromophore of GFP is formed from the precursory residues Ser65-Tyr66-Gly67 that are part of the proteins own polypeptide and located in the center of the β-barrel structure.[2,3]

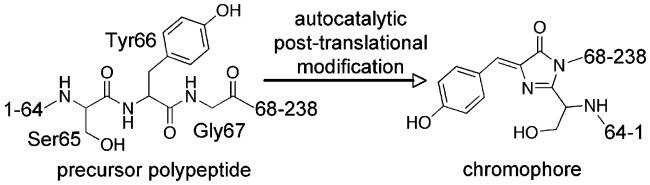

Figure 1. Chromophore formation in GFP.

Chromophore formation involves cyclization and dehydration of the main chain atoms to form an imidazolinone moiety, followed by oxidation of the Cα−Cβ bond of residue 66 (Figure 1). Thus, in the case of the GFP chromophore[4] with tyrosine at position 66, the imidazolinone formed from main chain atoms is brought into conjugation with the tyrosine-derived phenol. This overall process is surprisingly tolerant of modification of the key residues, providing protein engineers with the opportunity to alter the absorption and emission properties (i.e., the color) of the protein by changing either the covalent structure of the chromophore or the molecular environment of the cavity in which it resides. The most profound means of changing the color of a FP is to mutate residue 66 to an amino acid other than tyrosine. In the cases of the Y66F, Y66H, and Y66W substitutions, the altered chromphore retains an aromatic side chain that

* Corresponding authors. E-mail: robert.e.campbell@ualberta.ca; alex.brown@ualberta.ca.

10.1021/jp709900k CCC: $40.75
Published on Web 02/05/2008

is in conjugation with the main chain-derived imidazoline. The Y66H and Y66W substitutions are the key mutations of the popular variants known as blue FP (BFP)[5] and cyan FP (CFP),[6] respectively.

The engineering of the BFP and CFP variants is, arguably, one of the few examples in which researcher intuition has led to predictable and desirable changes in the properties of FP. For the most part, improvements in the properties of FPs are realized through the use of 'irrational' means. The most successful strategies for improving the properties of a FP are based on the high throughput screening of randomly mutated gene libraries. While the effectiveness of such approaches is clearly evident, the potential benefits of being able to perform computational screening or design of FP variants with improved spectral properties are substantial. Computational design holds the potential for sampling much more diverse regions of sequence space than is practical with in vitro methods to find variants with altered colors or improved brightness. Furthermore, computational methods could potentially be used to design variants with properties, such as extreme photostability, homogeneous fluorescence lifetime, and photoconversion, that are not readily amenable to high throughput in vitro identification. If we hope to ever apply computational methods to the design of FP variants with altered spectral properties, we first require a detailed understanding of the photochemical processes that result in fluorescence.

In light of the importance of FPs for biological applications and growing body of experimental data on the photochemistry of the chromophore, it is perhaps not surprising that computational modeling of FP chromophores with quantum-chemical methods has attracted considerable attention in recent years.[7−18] Photoexcitation of the GFP chromophore has been modeled in vacuum, in solvent, and in the context of the protein matrix using semiempirical,[7,18] single-reference[17] and multireference[10,11,13] *ab inito* methods. Time-dependent density functional theory (TD DFT)[19] also has been applied to the excited states of FP chromophores.[9,11−16,18,20] TD DFT calculations are often used in combined quantum mechanical/molecular mechanics (QM/MM) approaches,[13−15,18] which attempt to account for the effect of the protein environment and, in combination with molecular dynamics simulations,[18] to produce a realistic distribution of absorption energies of the chromophore in solvent and/or protein environment. Quantum-chemistry calculations have been used to interpret the experimental spectra of the FP chromophores and to assign certain absorption bands to particular ionization states or conformations of the chromophore.[16,20] The reported accuracy of TD DFT and other quantum-chemistry methods for reproducing experimental excitation energies of FP chromophores varies from one literature source to another. In some cases, the predicted excitation energies closely match the experimental data (errors <0.1 eV),[9] but in other cases, errors >0.4 eV are reported.[16] To the best of our knowledge, there has been no systematic analysis of the performance of various computational quantum chemistry methods across a series of different FP chromophores.

In this work, we report a self-consistent experimental data set of spectral properties for a structurally diverse series of FP chromophores and the computational prediction of the excitation energies using a variety of methods. To this end, we have genetically engineered and expressed versions of all previously reported GFP variants with structurally distinct chromophore structures and the previously unreported **Y66D** GFP variant. Absorption spectra of all chromophores in all chemically reasonable protonation states were measured for the denaturated proteins in order to remove the specific interactions of the chromophore with the protein environment. We chose not to include GFP variants with chromophores that were derived from unnatural amino acid analogues[21] or GFP homologues, including those such as DsRed,[22] with conjugation that extends beyond the imidazolinone ring. The experimental excitation energies and oscillator strengths were compared and correlated with computationally predicted values obtained using a variety of single-reference quantum-chemistry methods. This work has allowed us to critically evaluate the performance of these computational methods and to identify those methods that are best suited to the accurate prediction of the photochemical properties of FP chromophores.

2. Experimental Methods

Materials. All synthetic DNA oligonucleotides were purchased from Integrated DNA Technologies. A 1:1 v/v mixture of Taq polymerase (New England Biolabs) and Pfu polymerase (Fermentas) was used for all polymerase chain reaction (PCR) amplifications in 1× Pfu buffer supplemented with 2 mM $MgSO_4$. PCR products and products of restriction digest were routinely purified using GenCatch PCR Cleanup and Gel Extraction kits (Epoch Biolabs), respectively, according to the manufacturers protocols. T4 DNA ligase, Xho1 restriction enzyme, and EcoR1 restriction enzyme were purchased from New England Biolabs. Plasmid DNA was isolated from bacterial culture using GeneJET plasmid miniprep kit (Fermentas). The cDNA sequences for all GFP variants were confirmed by dye terminator cycle sequencing using the DYEnamic ET kit (Amersham Biosciences). Sequencing reactions were analyzed at the University of Alberta Molecular Biology Service Unit.

Site-Directed Mutagenesis. The proteins used as representative examples of the **Y66**, **Y66W**, and **Y66H** chromophores are EGFP (*Aequorea* GFP F64L, S65T),[23] ECFP (*Aequorea* GFP K26R, F64L, S65T, Y66W, N146I, M153T, V163A, N164H, H231L)[24] and EBFP (*Aequorea* GFP F64L, S65T, Y66H, Y145F),[25] respectively. For the other chromophores investigated, the representative examples are as follows: **Y66F** is EGFP T65A, Y66F; **Y66S** is EGFP T65A, Y66S; **Y66D** is EGFP T65A, Y66D; **Y66L** is EGFP T65A, Y66L; **Y66N** is EGFP T65A, Y66N; and **Y66T** is EGFP T65A, Y66T. Analogous EGFP Y66S, EGFP Y66D, and EGFP Y66N variants lacking the T65A mutation had no absorbance peak at wavelengths greater than 280 nm. Site-directed mutagenesis to change the identity of residues 65 and 66 was performed using an overlap extension PCR-based method.[26] The resulting PCR product encoding the full length EGFP gene (with the desired mutation) was digested with Xho1 and EcoR1 and ligated with similarly digested pBAD/HisB vector (Invitrogen). Electrocompetent *Escherichia coli* strain DH10B (Invitrogen) was transformed with the ligation reaction and plated on Luria−Bertani (LB)/agar plates supplemented with ampicillin (0.1 mg/mL) and L-arabinose (0.02%). Plates were incubated for 14 h at 37 °C prior to inspection. A single colony was used to inoculate 5 mL of LB liquid media supplemented with ampicillin. Plasmid DNA was isolated from the resulting culture following incubation for 24 h at 37 °C and 225 rpm.

Protein Purification. Protein purification procedure started with the inoculation of 0.5 L of LB media containing ampicillin (0.1 mg/mL) and L-arabinose (0.2%) with a single colony of *E. coli* LMG194 that had been transformed with the plasmid encoding the variant of interest. Cultures were grown overnight at 37 °C with shaking at 225 rpm. Cultures were cooled to 4 °C on ice, and cells were harvested by centrifugation (10 min,

8000 rpm). The cell pellet was resuspended in phosphate-buffered saline (PBS) (130 mM NaCl, 2.7 mM KCl, 10 mM phosphate, pH 7.4), and the cells were lysed by a single passage through a French pressure cell press (Thermo Electron). Insoluble cell debris was pelleted by centrifugation at 4 °C (20 min, 8000 rpm), and 0.7 mL of Ni-NTA resin (Qiagen) was added to the decanted supernatant. Following 30 min of gentle shaking at 4 °C, the suspended resin was loaded onto a 6 mL polypropylene column on a vacuum manifold. The resin was washed according to the manufacturers protocol and gravity eluted with 1.5 mL of a solution containing 200 mM imidazole (pH 7.0). Proteins were dialyzed overnight into 50 mM Tris buffer (pH 7.5) using a 7000 MWCO dialysis membrane (Pierce Snakeskin). To facilitate the oxidation steps required for formation of the **Y66L$_{ox1}$** and **Y66L$_{ox2}$** chromophores, EGFP T65A, Y66L was incubated in formate solution as previously described.[27] To attempt to form the **Y66L$_{ox1}$** chromophore, freshly purified EGFP T65A, Y66L was incubated in 2.0 M sodium formate, 100 mM cacodylic acid pH 6.5 for 2 days at 30 °C. To form the **Y66L$_{ox2}$** chromophore, EGFP T65A, Y66L was incubated in 2.0 M sodium formate, 100 mM Tris pH 8.0 for 1 week at 37 °C. Analogous incubations were attempted with EGFP T65A, Y66N and EGFP T65A, Y66T.

Spectroscopy. Solutions for protein denaturation were buffered at acidic pH (8 M urea, 30 mM citric acid, pH 2.3), neutral pH (8 M urea, 30 mM 3-(N-morpholino)propanesulfonic acid (MOPS), pH 7.0), or basic pH (8 M urea, 30 mM KH_2PO_4, pH 12.0). In order to form the anionic **Y66H$_{13}$** chromophore, the basic solution was adjusted to approximately pH 13. To prepare denatured protein for spectroscopic characterization, 20 μL of dialyzed protein was added to 180 μL of the buffered urea solution and then heated at 95 °C for 2 min. Absorbance spectra were recorded on a Beckman-Coulter DU800 spectrometer. Spectra were converted from an absorbance to an extinction coefficient (ϵ) scale by normalizing the absorption spectra of all folded proteins at 280 nm and then scaling such that ϵ for EGFP at 488 nm was 56 000 M^{-1} cm^{-1}.[1] Scaling by the ratio of peak intensities measured before and after denaturation provided the ϵ scale for the denatured proteins.

3. Computational Methods

Geometries of the model chromophores (Figures 2 and S2) were optimized using the hybrid Perdew, Burke, and Ernzerhof[28] exchange-correlation density functional (PBE0) with the standard 6-311++G(2df,2p) basis set for all atoms.[29] Geometry optimizations were performed in the gas phase and in water, using the IEF-PCM polarizable continuum model[30] as implemented in Gaussian03.[31] The properties of the excited states of the model chromophores were calculated with time-dependent density functional theory[19,32] (TD DFT) employing several hybrid and pure functionals, as well as with configuration interaction with single excitations[33] (CIS), and CIS with perturbative correction for double substitutions[34] [CIS(D)]. Excited-state calculations employed standard Pople-style double and triple-ζ basis sets, 6-31+G(d,p) and 6-311++G(2df,2p).[29] For TD DFT calculations, several hybrid and pure generalized gradient approximation (GGA), new-generation meta-GGA, and simple local spin density approximation (LSDA) functionals were employed. The GGA functionals included the exchange-correlation functional by Perdew, Burke, and Ernzerhof (PBE) and its hybrid version (PBE0),[28] Becke's 1988 exchange functional in combination with the correlation functional of Lee, Yang, and Parr (BLYP),[35] and Becke's three-parameter hybrid functional in combination with LYP correlation functional (B3LYP).[36] As well, pure and hybrid versions of Tao, Perdew, Staroverov, and Scuseria meta-GGA exchange-correlation functional (TPSS and TPSSh)[37] and LSDA Slater exchange functional[38] in combination with Vosko, Wilk, and Nusair correlation functional[39] (SVWN) were used. The solvent (water) effects on the excitation energies were treated using IEF-PCM model[40] equilibrated to the ground-state charge distribution in the molecule (nonequilibrium solvation of the excited state). All single-reference excited-state calculations, in both the gas phase and in PCM water, were carried out with Gaussian03.[31] Exploratory complete active space self-consistent field[41] (CASSCF) calculations were performed with the MOLPRO 2006.1 package.[42]

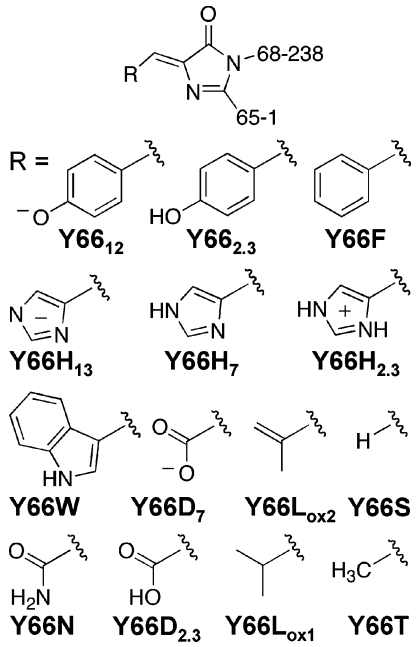

Figure 2. Chromophore structures investigated in this work. Each analogue is designated with a name that indicates the amino acid substitution at position 66 (tyrosine (Y) in wild-type) using standard single letter amino acid abbreviations. For structures with more than one possible ionization state, subscripts are used to designate the pH at which the experimental value of the absorption maxima was determined. The **Y66N**, **Y66D$_{2.3}$**, **Y66L$_{ox1}$**, and **Y66T** chromophores were attempted to be made but either did not form the expected structure and/or were not investigated in the computational aspect of this study.

4. Experimental Results

The chromophore structures investigated in this work (Figure 2) can be grouped into four categories: those which have previously been characterized in the context of both a folded GFP variant and in the denatured state (i.e., **Y66**, **Y66W**, **Y66H**); those which have been previously characterized in the context of the folded protein but not in the denatured state (i.e., **Y66F**, **Y66L**, and **Y66S**); those which have not been previously reported in any context and which did undergo all steps of chromophore maturation (i.e., **Y66D**); and those that have not been previously investigated and did not undergo all steps of chromophore maturation (i.e., **Y66N** and **Y66T**). The absorbance maxima, extinction coefficients, and estimated oscillator strengths are provided in Table 1, and the experimental absorbance spectra are provided in Supplementary Figure S1. In the following sections we provide brief descriptions of the experimental results for each category of chromophore.

Y66, Y66W, and Y66H. These chromophore structures correspond to those of wild-type GFP, the widely used CFP

TABLE 1: Experimentally Determined Absorption Maxima, Excitation Energies, Extinction Coefficients and Estimated Oscillator Strengths for Chromophores of Denatured Proteins

chromophore	λ_{max}, nm	ΔE, eV	ϵ, $mM^{-1}\,cm^{-1}$	f
Y66$_{12}$	446	2.78	45.0	0.7
Y66$_{2.3}$	378	3.28	34.5	0.65
Y66W	420	2.95	35.2	0.6
Y66F	350	3.54	20.9	0.45
Y66H$_7$	366	3.39	28.3	0.55
Y66H$_{2.3}$	352	3.52	16.7	0.4
Y66H$_{13}$	410	3.02	35.8	0.6
Y66D$_7$	322	3.85	13.0	~0.2
Y66S	328	3.78	24.0	0.5
Y66L$_{ox2}$	410	3.02	0.8	~0.15

variant, and the well-known BFP variant, respectively.[6,43] Accordingly, these chromophores have been the subject of thorough previous investigations with which our results are consistent.[5,44]

Y66F, Y66L, and Y66S. It has long been known that the **Y66F** variant is competent to undergo all steps of chromophore formation.[45] However, because UV-excitation is required to excite its fluorescence, this variant has not been used in biological studies.[1] In the folded state, the **Y66F** variant exhibited an absorbance maximum at 348 nm with a longer wavelength shoulder peaking at ~360 nm. Excitation at either peak resulted in fluorescence emission at 428 nm. Upon denaturation, a single absorbance peak with a maximum at 350 nm ($\epsilon = 20\,890\ cm^{-1}\,M^{-1}$, $f \sim 0.45$) was observed. Excitation at this peak resulted in no significant fluorescence.

The **Y66L** variant is nonfluorescent and not of practical use in biological labeling experiments. However, it has been previously investigated in efforts to understand the mechanism of GFP chromophore formation.[27] As previously reported, the protein required prolonged incubation in the presence of sodium formate to promote the oxidative steps of the chromophore maturation. Maturation of the **Y66L** chromophore requires two steps of oxidation: the first step producing **Y66L$_{ox1}$**, and the second step producing **Y66L$_{ox2}$** (Figure 2). Despite our best efforts, we were not able to identify conditions in which the **Y66L$_{ox1}$** state could be experimentally isolated, and so this chromophore is not included in our computational modeling. For **Y66L$_{ox2}$** there was no significant shift in the absorption peak at 410 nm ($\epsilon = 7922\ cm^{-1}\,M^{-1}$, $f \sim 0.15$) upon denaturation. The minimal wavelength shifts observed upon denaturation of the Y66F and Y66L variants suggest that there is little interaction between the chromophore and its surrounding cavity in the folded protein.

As with the **Y66L** variant, the **Y66S** variant is nonfluorescent but has previously investigated in studies aimed at elucidating the mechanism of GFP chromophore formation.[46] The absorption maximum was 350 nm in the folded protein and 328 nm ($\epsilon = 23\,969\ cm^{-1}\,M^{-1}$, $f \sim 0.5$) in the denaturated state. It has previously been determined that rather than undergoing the typical oxidation of the Cα−Cβ bond, the Y66S variant undergoes an elimination of water to form the terminal alkene-containing structure represented as Y66S in Figure 2.[46] This aromatic chromophore is similar in structure to the nonaromatic 4-methylidene-imidazole-5-one group which is derived from the sequence Ala-Ser-Gly and is observed in the active site of the enzyme histidine ammonia lyase.[47]

Y66D. This variant has, to the best of our knowledge, not been previously reported. The folded protein had a relatively strong absorbance peak at 350 nm which shifted to 322 nm ($\epsilon = 12\,966\ cm^{-1}\,M^{-1}$, $f \sim 0.2$) in the denatured protein. Using the pK_a of acrylate (4.25) as a guide, we would expect this chromophore to be in the anionic state (**Y66D$_7$**, Figure 2) at neutral pH. Indeed, at the lowest pH value attempted (2.3) we saw no significant shift in absorbance relative to the absorbance maxima measured at neutral and basic pH. As will be discussed later, similar energies for the neutral **Y66D$_{2.3}$** and anionic **Y66D$_7$** forms are predicted by the computational modeling.

TABLE 2: Excitation Energies of the Model Chromophores and Oscillator Strengths in the Gas Phase and Water Calculated with CIS

	CIS/6-31+G(d, p)				CIS/6-311++G(2df, 2p)			
	gas phase		PCM water		gas phase		PCM water	
chromophore	ΔE, eV	f	ΔE, eV	f	ΔE, eV	f	ΔE, eV	f
Y66$_{12}$	3.77	1.45	3.96	1.37	3.74	1.40	3.93	1.33
Y66$_{2.3}$	4.47	1.06	4.43	1.08	4.42	1.03	4.39	1.05
Y66W	4.29	0.88	4.22	0.93	4.25	0.86	4.19	0.91
Y66F	4.54	1.01	4.54	1.00	4.49	0.98	4.50	0.97
Y66H$_7$	4.59	0.79	4.52	0.83	4.54	0.78	4.48	0.81
Y66H$_{2.3}$	4.78	0.91	4.74	0.98	4.73	0.88	4.69	0.95
Y66H$_{13}$	4.07	1.23	4.25	1.20	4.03	1.21	4.21	1.18
Y66D$_7$	5.43	0.64	5.27	0.55	5.35	0.61	5.21	0.53
Y66S	5.58	0.47	5.55	0.42	5.50	0.43	5.48	0.39
Y66L$_{ox2}$	4.89	1.04	4.87	1.00	4.83	0.99	4.82	0.95

TABLE 3: Excitation Energies of the Model Chromophores in the Gas Phase and Water Calculated with CIS(D)

	CIS(D)/6-31+G(d,p)		CIS(D)/6-311++G(2df,2p)	
chromophore	gas phase	PCM water	gas phase	PCM water
Y66$_{12}$	2.82	3.02	2.82	3.00
Y66$_{2.3}$	4.00	3.89	3.91	3.81
Y66W	3.72	3.53	3.65	3.49
Y66F	4.21	4.15	4.10	4.05
Y66H$_7$	4.15	4.01	4.04	3.92
Y66H$_{2.3}$	4.22	4.21	4.10	4.10
Y66H$_{13}$	3.34	3.54	3.31	3.50
Y66D$_7$	4.84	4.48	4.72	4.37
Y66S	4.90	4.74	4.77	4.62
Y66L$_{ox2}$	4.42	4.34	4.30	4.23

Y66N and Y66T. If these variants had been competent to form the expected chromophore analogue structures, they would have produced the structures represented as **Y66N** and **Y66T** in Figure 2. To a first approximation, we expected these structures to have absorption maxima similar to that of **Y66D** and **Y66S**, respectively. Our experimental observation was that both variants exhibited weak absorbance between 300 and 350 nm in the folded state. Upon denaturation there was no significant absorbance at any wavelengths longer than the 280 nm peak attributed to tryptophan. These results are consistent with chromophore formation in these variants stalling after formation of the imidazolinone moiety or, more likely, forming a main-chain-derived heterocycle with an sp^3-hybridized ring nitrogen.[27] Upon denaturation the ring heterocycle is likely hydrolyzed to regenerate the original amino acid sequence.

5. Computational Results

Experimental results on the excitation energies and oscillator strengths for the unfolded FPs described in the previous section (Table 1) were compared against the results of quantum-chemistry calculations (Tables 2−4, S1−S6) for the model chromophores (Figure S2). The model chromophore structures (Figure S2) were obtained by terminating the chemical bonds connecting the chromophore to the rest of the polypeptide chain with hydrogen atoms (Figure 2). The lowest-energy conformations of the chromophores were used for the excited-state calculations. In all cases, the lowest-energy structure was the *Z*

TABLE 4: Excitation Energies of the Model Chromophores and Oscillator Strengths in the Gas Phase and Water Calculated with TD-PBE0

	TD-PBE0/6-31+G(d,p)				TD-PBE0/6-311++G(2df,2p)			
	gas phase		PCM water		gas phase		PCM water	
chromophore	ΔE, eV	f	ΔE, eV	f	ΔE, eV	f	ΔE, eV	f
Y66$_{12}$	3.15	0.93	3.04	1.06	3.12	0.91	3.02	1.04
Y66$_{2.3}$	3.61	0.72	3.41	0.85	3.57	0.70	3.38	0.83
Y66W	3.42	0.57	3.22	0.70	3.39	0.55	3.20	0.69
Y66F	3.72	0.62	3.56	0.71	3.68	0.60	3.53	0.69
Y66H$_7$	3.76	0.50	3.56	0.60	3.71	0.49	3.52	0.59
Y66H$_{2.3}$	3.76	0.42	3.65	0.55	3.72	0.40	3.61	0.54
Y66H$_{13}$	3.51	0.78	3.41	0.85	3.47	0.76	3.38	0.84
Y66D$_7$	2.93	0.01	4.02	0.27	2.94	0.01	3.99	0.26
Y66S	6.31	0.22	6.15	0.27	6.26	0.22	6.10	0.27
Y66L$_{ox2}$	3.93	0.46	3.77	0.51	3.89	0.44	3.73	0.50

stereoisomer (often referred to as cis) with respect to the bridging C═C bond which connects the imidozalinone moiety to the substituent R (Figure 2). This is the same stereoisomer as that experimentally observed in X-ray structures.[3,5,27,48]

We begin our discussion with the results obtained using configuration interaction with single excitations (CIS), the simplest *ab inito* method to treat electronically excited states.[33] According to CIS, the S_0 and S_1 excitation corresponds to the HOMO → LUMO transition of $\pi \rightarrow \pi^*$ type for most of the chromophores (Figure S2). The calculated excitation energies ΔE_{calc} of the model chromophores in the gas phase or PCM water (Table 2) are overestimated by 1.0−1.8 eV compared to the experimental values (Table 1). The largest errors are observed for the chromophores with nonaromatic substituents (**Y66D$_7$**, **Y66S**, **Y66L$_{ox2}$**). The basis set size does not have a considerable effect on the excitation energy: the triple-ζ results are just slightly lower than the double-ζ values (<0.1 eV difference). Inclusion of the solvent (water) effects with the PCM model leads to ≤0.2 eV shift (positive or negative) of the excitation energy, which is more pronounced for anionic species (**Y66$_{12}$**, **Y66H$_{13}$**, and **Y66D$_7$**).

The CIS method is known to overestimate grossly the excitation energies of organic molecules[49] and cannot be used for quantitative calculations of the excitation energies of the FP chromophores. On the other hand, quantum chemistry calculations are often used as a tool to distinguish between different forms of the FP chromophores (e.g., different conformations or ionization states) by comparing calculated and experimental energies.[12,16] Since the quantitative agreement between the calculated and experimental excitation energies is hard to achieve, it is important to be able to reproduce *relative* trends. Here, we used linear regression analysis to assess the correlation between calculated (ΔE_{calc}) and experimental (ΔE_{expt}) excitation energies. A linear function of the experimental excitation energy $\Delta E_{calc} = a + b\Delta E_{expt}$ was fitted with the least-squares procedure, and the results are reported in Table S8. Ideally, the intercept value a must be zero, and the slope $b = 1$. The plots of CIS-calculated versus experimental excitation energies presented in Figure 3 (a and b) demonstrate relatively good correlations with $R^2 > 0.85$. **Y66L$_{ox2}$** was excluded from the regression analysis since its ΔE_{calc} appears to be anomalously high: whereas the experimental excitation energies of **Y66L$_{ox2}$** and **Y66H$_{13}$** chromophores are practically the same, the calculated values for **Y66L$_{ox2}$** and **Y66H$_{13}$** differ by ~0.7 eV. Inclusion of **Y66L$_{ox2}$** into the regression decreases the R^2 to ~0.65. We address possible reasons behind this anomaly in some detail below. The basis set size does not have a significant effect on the correlation (Figures 3 and S3). Interestingly, the correlation is slightly worse when the solvation effects are included ($R^2 = 0.88$ in the gas phase versus 0.85 in PCM water). The value of the slope b is greater than 1 (Table S8), so the deviation between the calculated and experimental values increases as the excitation energy becomes higher.

The oscillator strengths predicted by CIS calculations tend to be significantly overestimated and do not seem to depend much on the inclusion of the solvation effects or the basis set used (Table 2). There is practically no correlation between the calculated and experimental oscillator strengths (Figure S4, Table S8).

The CIS results may be improved by accounting for nondynamic electron correlation using the CIS(D) approach, which

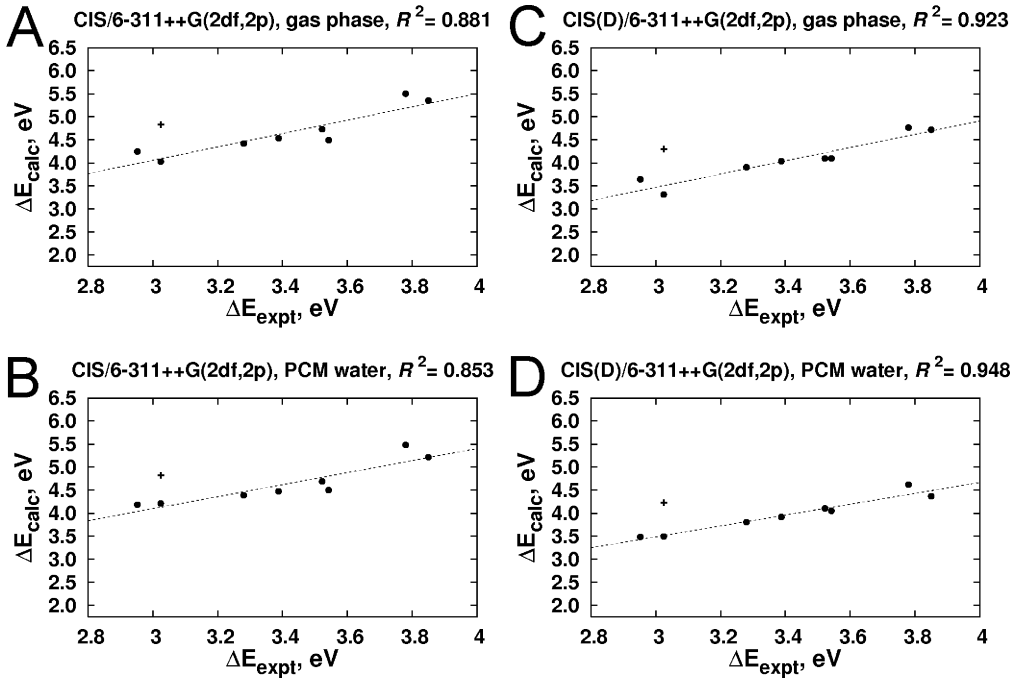

Figure 3. Correlation between calculated and experimental excitation energies: CIS/6-311++G(2df,2p) and CIS(D)/6-311++G(2df,2p) calculations. **Y66L$_{ox2}$** is excluded from the correlation and shown as "+".

perturbatively introduces effects of double excitations.[34] The excitation energies calculated with CIS(D) (Table 3) are closer to the experimental data (Table 1), although they are still 0.2−0.5 eV higher. For aliphatically substituted chromophores **Y66S** and **Y66L$_{ox2}$** the error is even higher (0.8−1.3 eV). Limited inclusion of the electron correlation with the CIS(D) method leads to improved correlation of ΔE_{calc} with experimental values (Figure 3, c and d), $R^2 > 0.92$. Although the excitation energy of **Y66L$_{ox2}$** is somewhat improved at the CIS(D) level, it still does not belong to the ΔE_{calc} versus ΔE_{expt} correlation observed for the other chromophores (Figure 3, c and d), and therefore it has been excluded from the linear regression analysis. Unlike CIS (Figure 3, a and b), for CIS(D) the inclusion of the solvation effects improves correlation ($R^2 = 0.95$ in water versus 0.92 in the gas phase). The ΔE_{calc} values calculated with CIS(D) are slightly more sensitive to the increase of the basis set size from double to triple-ζ (> 0.1 eV difference for some chromophores), but the quality of the correlation ΔE_{calc} versus ΔE_{expt} does not change.

Currently, time-dependent density functional theory (TD DFT) is the most widely used quantum-chemical method for calculation of the excited states for medium-size organic molecules.[49] Since density functional theory includes electron correlation, TD DFT can provide significantly better results for the excitation energies and oscillator strengths than CIS, at similar computational cost. A number of TD DFT calculations of various FP chromophores have been reported in the literature.[12,14,16] Here, we performed TD DFT calculations using a number of commonly used hybrid and pure density functionals, including generalized gradient approximation (GGA) functionals such as PBE, BLYP (pure), PBE0 and B3LYP (hybrid), meta-GGA TPSS (pure) and TPSSh (hybrid) functionals, and the local-density approximation SVWN functional. The excitation energies and oscillator strengths calculated with TD-PBE0 are presented in Table 4, and the data obtained with the other functionals are presented in the Supporting Information section (Tables S1−S6).

While the TD DFT calculations using hybrid functionals suggest a very similar $\pi \rightarrow \pi^*$ nature of the $S_0 \rightarrow S_1$ transition in the model chromophores, in agreement with the CIS calculations (Figure S2), admixtures of transitions involving other orbitals may be observed when pure functionals are used, as has been reported previously in the literature.[9,16] The excitation energies calculated with TD-PBE0 (Table 4) are generally overestimated by 0.3−0.5 eV compared to the experimental data (Table 1), in agreement with the literature.[16] When the solvent effects are included, the excitation energy of **Y66F** is reproduced quantitatively (3.54 versus 3.53 eV), probably because of mutual cancelation of errors. On the other hand, there are significant errors in the calculated excitation energies for several nonaromatic chromophores. The excitation energy of **Y66S** is overestimated by >2.4 eV, irrespective to inclusion of solvent effects or the basis set used. Interestingly, the **Y66D$_7$** excitation energy calculated in the gas phase is underestimated by ~1.1 eV, but when the solvation is included, the predicted excitation energy is within 0.2 eV from the experimental value (3.90 eV versus 3.85 eV experimental). No such dramatic influence of the solvent effects was observed in the case of CIS and CIS(D) calculations (Tables 2 and 3, Figure 3). Similarly to CIS and CIS(D), TD-PBE0 significantly overestimates the excitation energy for **Y66L$_{ox2}$** (by 0.7−0.9 eV).

Calculations with the popular B3LYP functional yield very similar results (Table S2), and even smaller errors for some chromophores. However, TD-B3LYP errors in the case of **Y66S**, **Y66L$_{ox2}$** and **Y66D$_7$** chromophores calculated in the gas phase are unacceptably large, similarly to TD-PBE0. There is very little basis set dependence for the excitation energies calculated with TD DFT. Pure functionals PBE and BLYP generally yield excitation energies which are underestimated by 0.1−0.3 eV (Tables S1 and S3), while the errors in the case of the **Y66S**, **Y66L$_{ox2}$**, and **Y66D$_7$** are similar to the errors observed for the hybrid functionals. Results obtained with meta-GGA hybrid TPSSh functional (Table S4) are very similar to GGA hybrid functionals, and its pure version, TPSS (Table S5), performs similarly to the pure GGA functionals used in this work. Moreover, even the simple local density approximation SVWN functional gives results very similar to more advanced GGA and meta-GGA functionals (Table S6).

Excitation energies calculated with TD DFT correlate relatively well with the experimental data, if the **Y66S** and **Y66L$_{ox2}$** points are excluded (**Y66D$_7$** gas-phase excitation energy is also excluded). Figure 4 (a and b) shows the correlations for the excitation energies calculated with TD-PBE0; the correlations for the other functionals look very similar (e.g., Figure S5). The correlation parameters between TD DFT excitation energies obtained in this work and the experimental data are collected in Table S7. Although hybrid functionals yield larger errors for the absolute excitation energies (see above), ΔE_{calc} obtained with these functionals correlates better with the experimental values, while the pure functionals give more scattered data (e.g., $R^2 = 0.91$ and 0.83 for B3LYP and BLYP functionals, respectively). Increase of the basis set size and/or inclusion of the solvation effects systematically improves the correlation (Table S7). The best correlation is obtained with TD-PBE0/6-311++G(2df,2p) in PCM water ($R^2 = 0.92$). TD DFT correlations have slopes less than 1, unlike CIS and CIS(D) correlations where it is larger than 1. Good correlations of the TD-PBE0 calculated excitation energies with the experimental data have been previously reported for coumarins and chromones in various solvents.[50]

Similarly to CIS, TD DFT overestimates values of the oscillator strengths for the $S_0 \rightarrow S_1$ transitions for the model chromophores (Tables 4, S1−S6). However, the oscillator strengths calculated with TD DFT correlate much better with the experimental values, as can be seen from Table S8, which presents the correlation parameters for calculated versus experimental oscillator strengths, and Figures 4 (c and d) and S5 (b). Inclusion of the solvation effects is crucial for oscillator strength calculations: R^2 values are ~0.6 for the gas-phase calculations, whereas in solvent R^2 is always larger than 0.83. This is partially related to the fact that the correlations for PCM-calculated values include **Y66D$_7$**, which is excluded from the correlation for the gas-phase data. Interestingly, the best correlations are obtained for the calculations using pure functionals BLYP and SVWN ($R^2 = 0.91$).

TD DFT performs almost on par with CIS(D) in terms of the correlation between the calculated excitation energies and the experimental data, although the best correlation is observed in the case of PCM CIS(D)/6-311++G(2df,2p) model chemistry ($R^2 = 0.95$), whereas TD DFT yields slightly more scattered data [$R^2 = 0.92$ with PCM TD-PBE0/6-311++G(2df,2p)]. At the same time, TD DFT does not seem to work for **Y66S** chromophore, as well as **Y66D$_7$** chromophore in gas phase. Also, **Y66L$_{ox2}$** chromophore seems to be problematic for all methods used in this work. We discuss these 'difficult cases' in more detail below.

CIS calculations (in the gas phase or including solvation effects) describe the photoexcitation of **Y66D$_7$** to involve

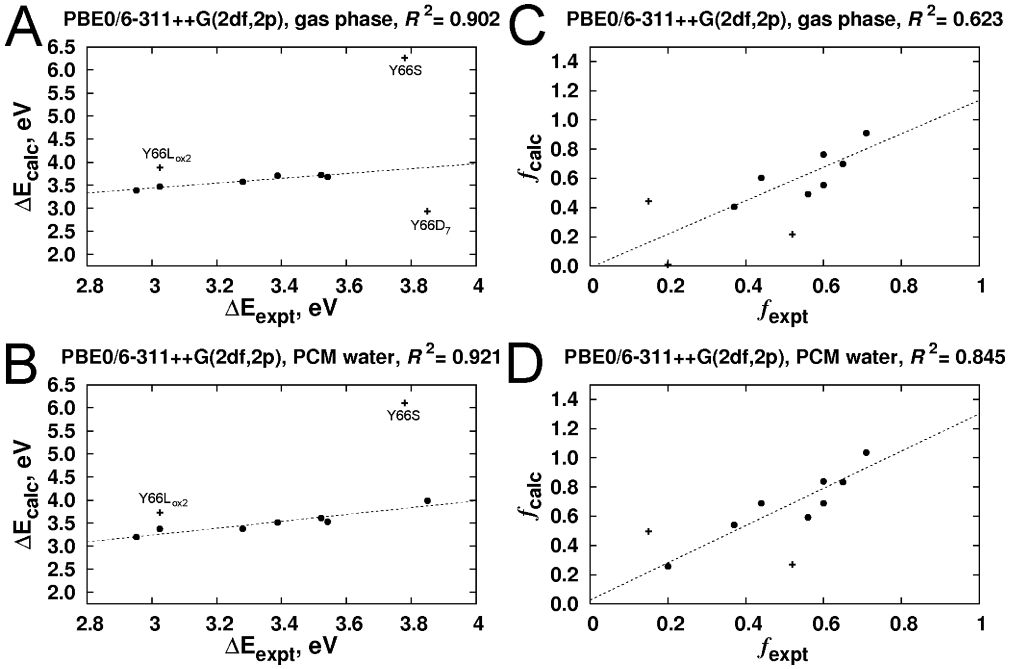

Figure 4. Correlation between calculated and experimental excitation energies (A and B) and oscillator strengths (C and D) of the model chromophores: PBE0/6-311++G(2df,2p) calculations. **Y66S**, **Y66L$_{ox2}$**, and **Y66D**$_7$ are excluded from the correlation in the case of gas-phase calculations, while **Y66S** and **Y66L$_{ox2}$** are excluded in the case of calculations in water. The excluded points are shown as "+".

transition from the π-type orbital which has a bonding character in respect to the bridging C═C and the imidozolinone N═C double bonds to a π*-type orbital with antibonding character with respect to the bridging C═C bond and bonding character in respect to the C−C bond of the imidozoline ring (Figure S6). TD−DFT calculations in solvent provide essentially the same picture of the $S_0 \rightarrow S_1$ transition in **Y66D$_7$**. On the other hand, in the gas phase, TD−DFT predicts the photoexcitaion to happen from the electron-rich π-system of the carboxylate moiety, leading to significant lowering of the excitation energy. Thus, inclusion of the solvation effects is essential not only for quantitative but also correct qualitative description of the photoexcitation in **Y66D$_7$**. We also note that the neutral **Y66D$_{2.3}$** model chromophore has similar excitation energy and oscillator strength calculated with TD-PBE0/6-311++G(2df,2p) in PCM water (ΔE = 3.71 eV and f = 0.23) as the anionic **Y66D$_7$** form (ΔE = 3.99 eV and f = 0.23), which agrees well with the experimental observation that the chromophore spectrum does not significantly change in neutral or acidic environments (Figure S1). On the other hand, in the case of the neutral **Y66D$_{2.3}$** form, the solvent does not have the same dramatic effect as for the anionic **Y66D$_7$** form: the excitation energy and oscillator strength do not change significantly going from the gas phase to water solution [ΔE = 3.87 versus 3.71 eV and f = 0.23 versus 0.22 at the TD-PBE0/6-311++G(2df,2p) level].

TD DFT greatly overestimates the excitation energy of the **Y66S** model chromophore, which has to be excluded from the correlation (Figure 4, a and b). Although CIS calculations predict the **Y66S** excitation energy which correlates relatively well with the experimental results, the correlation significantly improves upon its exclusion (R^2 = 0.85 versus 0.91, Table S8). Inclusion of the perturbative double excitations at the CIS(D) level significantly improves the correlation between the calculated and experimental excitation energies when **Y66S** is included in the correlation (R^2 = 0.95 and 0.96, with and without **Y66S**, respectively, Table S8). We performed exploratory complete active space self-consistent field[41,42] (CASSCF) calculations for the first excited-state of the **Y66S** model chromophore, with an (8,8) active space which included the four highest occupied and four lowest unoccupied π-type orbitals. The analysis of the configuration interaction (CI) expansion coefficients of the optimized CASSCF wavefunction for the first bright excited-state of **Y66S** chromophore shows that in addition to the singly excited configuration which corresponds to simple HOMO → LUMO transition (~82% contribution to the excited-state wavefunction):

$$\ldots (\pi)^2(\pi)^2(\pi)^2(\pi)^1(\pi^*)^1$$

an additional doubly excited configuration is the second important configuration (~4% contribution, the other configurations have <1% contribution):

$$\ldots (\pi)^2(\pi)^2(\pi)^1(\pi)^1(\pi^*)^2$$

On the other hand, the first bright excited-state for most of the other chromophores, e.g., **Y66$_{12}$** [calculated with CASSCF-(10, 10)], have only one, singly excited dominant configuration (~85% contribution, the other configurations have <1% contribution):

$$\ldots (\pi)^2(\pi)^2(\pi)^2(\pi)^2(\pi)^1(\pi^*)^1$$

The TD DFT approach is known to fail in the case of doubly excited states, which are often observed for polyenes, such as butadiene and its analogues.[51] Unlike most of the chromophores studied in this work, **Y66S** does not have aromatic character, but possesses conjugated C═C and C═N double bond motif similar to polyenes. The failure of TD DFT and (to a lesser extent) CIS methods in the case of the **Y66S** chromophore can be rationalized by significant contribution of the doubly excited configuration in the excited-state wavefunction. Better performance of CIS in this case is probably related to mutual cancelation of errors.

The CASSCF calculations for the **Y66L$_{ox2}$** chromophore using (8,8) active space also demonstrate significant contribution of a doubly excited configuration to the first bright excited-

state wavefunction, similarly to **Y66S** chromophore (~7% contribution):

$$\ldots (\pi)^2(\pi)^2(\pi)^1(\pi)^1(\pi^*)^2$$

However, an additional singly excited configuration which involves excitation from a low-lying π-orbital is also important in the case of **Y66L$_{ox2}$**(~4% contribution):

$$\ldots (\pi)^2(\pi)^1(\pi)^2(\pi)^2(\pi^*)^1$$

Apparently, the complex character of the bright excited-state of the **Y66L$_{ox2}$** chromophore is responsible for the poor description of the photoexcitation not only with the TD DFT method but also with CIS and CIS(D) methods. On the other hand, complex interplay between various singly and doubly excited components of the excited-state wavefunction probably leads to partial cancelation of errors, since the predicted **Y66L$_{ox2}$** excitation energy seems to correlate better with the experimental value, compared to **Y66S** excitation energy calculated with TD DFT (Figure 4, a and b).

To conclude, quantum-chemistry calculations with CIS(D) and TD DFT methods provide good qualitative and semiquantitative description of the relative differences between the excitation energies of the GFP-analogue chromophores, except certain nonaromatic chromophores. Comparison between the experimental and calculated excitation energies can be used to distinguish between different chromophores and/or various forms of one chromophore (different ionization states or conformations). However, if the experimental excitation energies differ by <0.1 eV, the calculated excitation energies cannot be used to distinguish between the chrtomophores/chromophore forms since the quantum-chemistry calculations are not reliable within <0.1 eV energy difference. For instance, the excitation energy of **Y66D$_7$** is 0.07 eV higher than **Y66S** (3.85 and 3.78 eV, respectively); on the other hand, PCM CIS(D)/6-311++G-(2df,2p) calculations predict **Y66S** excitation energy 0.25 eV higher than **Y66D$_7$** (4.62 and 4.37 eV, respectively).

Finally, we remark that the experimental spectra obtained in this work (Figure S1) are measured not in pure water but in 8 M urea solution, while the polarizable continuum model used here was parametrized for pure water. However, since all the measurements were done at similar conditions, the differences between water and water−urea solution apparently contribute to the systematic error in calculated spectral properties and do not impair the correlation between the calculated and experimental values. Also, good correlations between the excitation energies calculated using PCM solvation model and the experimental data suggest that the specific solvation of the chromophores does not significantly alter their excitation energies, in agreement with the literature.[16]

Conclusions

In this contribution, we presented a combined experimental and computational study of the photoexcitation of GFP analogue chromophores in order to systematically assess the performance of various single-reference quantum-chemistry methods for excited-state calculations. Seven GFP variants were expressed and purified, including the novel Y66D mutant. The excitation energies and oscillator strengths for 10 chemically distinct chromophores, inclusive of various ionization states of the same chromophore, were obtained by measuring absorption spectra of denaturated proteins. This experimental data set was compared against the excitation energies and oscillator strengths calculated for model chromophore structures using several quantum-chemistry methods, such as CIS, CIS(D), and TD DFT with various pure and hybrid density functionals.

The methods used are not able to quantitatively reproduce either the excitation energies or the oscillator strengths for the chromophores studied. It should be stressed that occasional quantitative agreements between the calculated and experimental values for certain chromophore/method combination is most likely because of cancelation of errors, and the application of the same method to another chromophore may lead to significant errors in the calculated excitation energies. However, a good linear correlation between the calculated and experimental excitation energies was observed, with the exception of several nonaromatic chromophores. The observed correlation suggests that the computational methods correctly describe the nature of the bright excited states of the GFP analogue chromophores. Since the quantum-chemistry calculations adequately reproduce relative trends, the calculated excitation energies can be used for spectral assignments of the experimental absorption spectra to different chromophore structures, if the experimental excitation energies differ by >0.1 eV.

The inclusion of the solvent effects using a polarizable continuum model is very important to obtain good correlation between the calculated and experimental excitation energies, especially in the case of TD DFT calculations. Calculated excitation energies for nonaromatic chromophores **Y66S** and **Y66L$_{ox2}$** are significantly overestimated with the CIS and TD DFT methods and do not correlate well with the experimental data. Preliminary CASSCF calculations suggest that the excited states of these chromophores have considerable contribution of doubly excited configurations, which cannot be treated with methods such as standard TD DFT or CIS. Thus, caution must be exercised when dealing with the excited states of nonaromatic chromophores; fortunately, most of the practically important GFP analogue proteins have aromatic chromophores.

To the best of our knowledge, detailed benchmarking of the calculated oscillator strengths for the GFP analogue chromophores has not been reported in the literature. In this work, we found that the oscillator strength values calculated with TD DFT, although overestimated, correlate well with the experimental data. Inclusion of solvent effects is essential for the oscillator strength calculations, and pure density functionals generally give better results. At the same time, oscillator strengths obtained with the CIS method do not correlate with the experiment.

We believe that the experimental data set reported in this work will be useful for further benchmarking of methods for excited-state calculations (e.g., semiempirical methods) and/or solvation models. The systematic study of spectral properties across the series of GFP analogue chromophores shows that the excited states of these relatively large and complex molecules can be reliably modeled with medium-level quantum chemistry methods. These methods may also be expected to perform well for the calculation of the emission spectra and, in QM/MM context, for the excited-state dynamics studies of FPs.

Acknowledgment. This work was supported by grants from the Natural Sciences and Engineering Research Council of Canada (NSERC) (to both R.E.C. and A.B.), an Alberta Ingenuity New Faculty Award (to R.E.C.), and the Canadian Foundation for Innovation. Q.K.T. was supported by an Alberta Ingenuity postdoctoral fellowship. The authors thank the University of Alberta and the Canadian Foundation for Innovation for infrastructure support. R.E.C. holds a Canada Research Chair in Bioanalytical Chemistry.

Supporting Information Available: Experimental absorbance spectra, model chromophore structures and shapes of the

molecular orbitals involved in the $S_0 \rightarrow S_1$ transition of the model chromophores, calculated excitation energies and oscillator strengths, their correlations with the experimental data, summary of the correlation parameters between the calculated and experimental data, and Cartesian coordinates for the model chromophores. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Tsien, R. Y. *Annu. Rev. Biochem.* **1998**, *67*, 509.

(2) Yang, F.; Moss, L. G.; Phillips, G. N., Jr. *Nat. Biotechnol.* **1996**, *14*, 1246.

(3) Ormo, M.; Cubitt, A. B.; Kallio, K.; Gross, L. A.; Tsien, R. Y.; Remington, S. J. *Science* **1996**, *273*, 1392.

(4) Shimomura, O. *FEBS Lett.* **1979**, *104*, 220; Cody, C. W.; Prasher, D. C.; Westler, W. M.; Prendergast, F. G.; Ward, W. W. *Biochemistry* **1993**, *32*, 1212.

(5) Wachter, R. M.; King, B. A.; Heim, R.; Kallio, K.; Tsien, R. Y.; Boxer, S. G.; Remington, S. J. *Biochemistry* **1997**, *36*, 9759.

(6) Heim, R.; Prasher, D. C.; Tsien, R. Y. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 12501.

(7) Voityuk, A. A.; Michel-Beyerle, M.-E.; Rosch, N. *Chem. Phys. Lett.* **1997**, *272*, 162; Weber, W.; Helms, V.; McCammon, J. A.; Langhoff, P. W. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 6177; Follenius-Wund, A.; Bourotte, M.; Schmitt, M.; Iyice, F.; Lami, H.; Bourguignon, J.-J.; Haiech, J.; Pigault, C. *Biophys. J.* **2003**, *85*, 1839; Toniolo, A.; Granucci, G.; Martinez, T. J. *J. Phys. Chem. A* **2003**, *107*, 3822; Toniolo, A.; Olsen, S.; Manohar, L.; Martinez, T. J. *Faraday Discuss.* **2004**, *127*, 149.

(8) Tozzini, V.; Nifosi, R. *J. Phys. Chem. B* **2001**, *105*, 5797; Wang, S.; Smith, S. C. *J. Phys. Chem. B* **2006**, *110*, 5084; Helms, V. *Curr. Opin. Struct. Biol.* **2002**, *12*, 169.

(9) Marques, M. A. L.; López, X.; Varsano, D.; Castro, A.; Rubio, A. *Phys. Rev. Lett.* **2003**, *90*, 258101.

(10) Martin, M. E.; Negri, F.; Olivucci, M. *J. Am. Chem. Soc.* **2004**, *126*, 5452; Altoe, P.; Bernardi, F.; Garavelli, M.; Orlandi, G.; Negri, F. *J. Am. Chem. Soc.* **2005**, *127*, 3952; Olsen, S.; Smith, S. C. *J. Am. Chem. Soc.* **2007**, *129*, 2054.

(11) Vendrell, O.; Gelabert, R.; Moreno, M.; Lluch, J. M. *Chem. Phys. Lett.* **2004**, *396*, 202.

(12) Lopez, X.; Marques, M. A. L.; Castro, A.; Rubio, A. *J. Am. Chem. Soc.* **2005**, *127*, 12329.

(13) Amat, P.; Granucci, G.; Buda, F.; Persico, M.; Tozzini, V. *J. Phys. Chem. B* **2006**, *110*, 9348.

(14) Grigorenko, B.; Savitsky, A.; Topol, I.; Burt, S.; Nemukhin, A. *J. Phys. Chem. B* **2006**, *110*, 18635.

(15) Grigorenko, B.; Savitsky, A.; Topol, I.; Burt, S.; Nemukhin, A. *Chem. Phys. Lett.* **2006**, *424*, 184; Patnaik, S. S.; Trohalaki, S.; Naik, R. R.; Stone, M. O.; Pachter, R. *Biopolymers* **2007**, *85*, 253.

(16) Nemukhin, A. V.; Topol, I. A.; Burt, S. K. *J. Chem. Theory Comput.* **2006**, *2*, 292.

(17) Mochizuki, Y.; Nakano, T.; Amari, S.; Ishikawa, T.; Tanaka, K.; Sakurai, M.; Tanaka, S. *Chem. Phys. Lett.* **2007**, *433*, 360.

(18) Schäfer, Lars V.; Groenhof, G.; Klingen, Astrid R.; Ullmann, G. M.; Boggio-Pasqua, M.; Robb, M. A.; Grubmüller, H. *Angew. Chem., Int. Ed.* **2007**, *46*, 530.

(19) Casida, M. E.; Jamorski, C.; Casida, K. C.; Salahub, D. R. *J. Chem. Phys.* **1998**, *108*, 4439.

(20) Olsen, S.; Prescott, M.; Wilmann, P.; Battad, J.; Rossjohn, J.; Smith, S. C. *Chem. Phys. Lett.* **2006**, *420*, 507.

(21) Wang, J. W.; Wong, A. M.; Flores, J.; Vosshall, L. B.; Axel, R. *Cell* **2003**, *112*, 271; Hyun Bae, J.; Rubini, M.; Jung, G.; Wiegand, G.; Seifert, M. H.; Azim, M. K.; Kim, J. S.; Zumbusch, A.; Holak, T. A.; Moroder, L.; Huber, R.; Budisa, N. *J. Mol. Biol.* **2003**, *328*, 1071; Budisa, N.; Pal, P. P. *Biol. Chem.* **2004**, *385*, 893; Budisa, N.; Pal, P. P.; Alefelder, S.; Birle, P.; Krywcun, T.; Rubini, M.; Wenger, W.; Bae, J. H.; Steiner, T. *Biol. Chem.* **2004**, *385*, 191; Pal, P. P.; Bae, J. H.; Azim, M. K.; Hess, P.; Friedrich, R.; Huber, R.; Moroder, L.; Budisa, N. *Biochemistry* **2005**, *44*, 3663.

(22) Matz, M. V.; Fradkov, A. F.; Labas, Y. A.; Savitsky, A. P.; Zaraisky, A. G.; Markelov, M. L.; Lukyanov, S. A. *Nat. Biotechnol.* **1999**, *17*, 969; Baird, G. S.; Zacharias, D. A.; Tsien, R. Y. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 11984; Gross, L. A.; Baird, G. S.; Hoffman, R. C.; Baldridge, K. K.; Tsien, R. Y. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 11990; Heikal, A. A.; Hess, S. T.; Baird, G. S.; Tsien, R. Y.; Webb, W. W. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 11996; Yarbrough, D.; Wachter, R. M.; Kallio, K.; Matz, M. V.; Remington, S. J. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 462.

(23) Patterson, G. H.; Knobel, S. M.; Sharif, W. D.; Kain, S. R.; Piston, D. W. *Biophys. J.* **1997**, *73*, 2782.

(24) Miyawaki, A.; Llopis, J.; Heim, R.; McCaffery, J. M.; Adams, J. A.; Ikura, M.; Tsien, R. Y. *Nature* **1997**, *388*, 882.

(25) Yang, T. T.; Sinai, P.; Green, G.; Kitts, P. A.; Chen, Y. T.; Lybarger, L.; Chervenak, R.; Patterson, G. H.; Piston, D. W.; Kain, S. R. *J. Biol. Chem.* **1998**, *273*, 8212.

(26) Ho, S. N.; Hunt, H. D.; Horton, R. M.; Pullen, J. K.; Pease, L. R. *Gene* **1989**, *77*, 51.

(27) Rosenow, M. A.; Patel, H. N.; Wachter, R. M. *Biochemistry* **2005**, *44*, 8303.

(28) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865.

(29) Hehre, W. J.; Radom, L.; Schleyer, P. v. R.; Pople, J. A. *Ab Initio Molecular Orbital Theory*; John Wiley and Sons: New York, 1985.

(30) Cances, E.; Mennucci, B.; Tomasi, J. *J. Chem. Phys.* **1997**, *107*, 3032; Cossi, M.; Scalmani, G.; Rega, N.; Barone, V. *J. Chem. Phys.* **2002**, *117*, 43.

(31) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr., T. V.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*, Revision B.05; Gaussian Inc., Pittsburgh, PA, 2003.

(32) Stratmann, R. E.; Scuseria, G. E.; Frisch, M. J. *J. Chem. Phys.* **1998**, *109*, 8218.

(33) Foresman, J. B.; Head-Gordon, M.; Pople, J. A.; Frisch, M. J. *J. Phys. Chem.* **1992**, *96*, 135.

(34) Head-Gordon, M.; Rico, R. J.; Oumi, M.; Lee, T. J. *Chem. Phys. Lett.* **1994**, *219*, 21.

(35) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098; Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.

(36) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623.

(37) Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. *Phys. Rev. Lett.* **2003**, *91*, 146401; Staroverov, V. N.; Scuseria, G. E.; Tao, J.; Perdew, J. P. *J. Chem. Phys.* **2003**, *119*, 12129.

(38) Hohenberg, P.; Kohn, W. *Phys. Rev.* **1964**, B864.

(39) Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, *58*, 1200.

(40) Cammi, R.; Mennucci, B.; Tomasi, J. *J. Phys. Chem. A* **2000**, *104*, 5631; Cossi, M.; Barone, V. *J. Chem. Phys.* **2001**, *115*, 4708.

(41) Werner, H. J.; Knowles, P. J. *J. Chem. Phys.* **1985**, *82*, 5053.

(42) Werner, H.-J.; Knowles, P. J.; Lindh, R.; Schütz, M.; Celani, P. ; Korona, T.; Rauhut, G.; Amos, R. D.; Bernhardsson, A.; Berning, A.; Cooper, D. L.; Deegan, M. J. O.; Dobbyn, A. J.; Eckert, F.; Hampel, C.; Hetzer, G.; Lloyd, A. W.; McNicholas, S. J.; Meyer, W.; Mura, M. E.; Nicklass, A.; Palmieri, P.; Pitzer. R.; Schumann, U.; Stoll, H.; Stone, A. J.; Tarroni, R.; Thorsteinsson, T. Molpro version 2006.1, a package of ab initio programs, see http://www.molpro.net.

(43) Heim, R.; Tsien, R. Y. *Curr. Biol.* **1996**, *6*, 178.

(44) Ward, W. W.; Cody, C. W.; Hart, R. C.; Cormier, M. J. *Photochem. Photobiol.* **1980**, *31*, 611; Ward, W. W.; Bokman, S. H. *Biochemistry* **1982**, *21*, 4535.

(45) Cubitt, A. B.; Heim, R.; Adams, S. R.; Boyd, A. E.; Gross, L. A.; Tsien, R. Y. *Trends Biochem. Sci.* **1995**, *20*, 448.

(46) Barondeau, D. P.; Kassmann, C. J.; Tainer, J. A.; Getzoff, E. D. *Biochemistry* **2005**, *44*, 1960.

(47) Schwede, T. F.; Retey, J.; Schulz, G. E. *Biochemistry* **1999**, *38*, 5355.

(48) Bae, J. H.; Rubini, M.; Jung, G.; Wiegand, G.; Seifert, M. H. J.; Azim, M. K.; Kim, J. S.; Zumbusch, A.; Holak, T. A.; Moroder, L.; Huber, R.; Budisa, N. *J. Mol. Biol.* **2003**, *328*, 1071.

(49) Dreuw, A.; Head-Gordon, M. *Chem. Rev.* **2005**, *105*, 4009.

(50) Preat, J.; Jacquemin, D.; Wathelet, V.; Andre, J.-M.; Perpete, E. A. *J. Phys. Chem. A* **2006**, *110*, 8144.

(51) Lappe, J.; Cave, R. J. *J. Phys. Chem. A* **2000**, *104*, 2294; Tozer, D. J.; Amos, R. D.; Handy, N. C.; Roos, B. O.; Serrano-Andres, L. *Mol. Phys.* **1999**, *97*, 859.