SCIENTIFIC REPORTS

OPEN

Long- and Short-Range Electrostatic Fields in GFP Mutants: Implications for Spectral Tuning

M. Drobizhev[1], P. R. Callis[2], R. Nifosì[3], G. Wicks[1], C. Stoltzfus[1], L. Barnett[4], T. E. Hughes[4], P. Sullivan[2] & A. Rebane[1,5]

The majority of protein functions are governed by their internal local electrostatics. Quantitative information about these interactions can shed light on how proteins work and allow for improving/altering their performance. Green fluorescent protein (GFP) and its mutation variants provide unique optical windows for interrogation of internal electric fields, thanks to the intrinsic fluorophore group formed inside them. Here we use an all-optical method, based on the independent measurements of transition frequency and one- and two-photon absorption cross sections in a number of GFP mutants to evaluate these internal electric fields. Two physical models based on the quadratic Stark effect, either with or without taking into account structural (bond-length) changes of the chromophore in varying field, allow us to separately evaluate the long-range and the total effective (short- and long-range) fields. Both types of the field quantitatively agree with the results of independent molecular dynamic simulations, justifying our method of measurement.

Electric fields inside proteins play a central role in enzymatic function[1,2], firing action potentials in neurons[3,4], color vision[5], etc. In GFP and its mutants, a variety of spectral hues is usually attributed to different electrostatic environments of their chromophore. By mutating the local protein surrounding, while keeping the structure of the chromophore unchanged, one can tune the optical transition frequency within an impressive spectral range of ~65 nm[6,7], see Fig. 1. Structural analysis[7] and quantum simulations[8–11] suggest that the main part of this tuning is due to variations of electrostatic fields around the chromophore. Therefore, on one hand, one can use GFP and its mutants as model systems to investigate internal electric fields, especially because the structure of many of them is known, and on the other hand, this knowledge could lead to the creation of new fluorescent proteins (FPs) with improved optical properties.

Classical Stark shift methods measure internal fields in proteins by using either UV-vis (electronic) or infrared (vibrational) absorption spectroscopy[2,12,13]. The development of alternative approaches, based on a combination of one- and two-photon excited fluorescence, would allow evaluating these fields in biologically intact living cells and tissues and in the high-resolution fluorescence microscopy conditions.

The one-photon absorption (1PA) spectrum of a chromophore in a weak uniform electric field E experiences a spectral shift $\Delta\nu$ ($\Delta\nu = \nu - \nu_0$, where ν and ν_0 are the transition frequencies in protein and in vacuum, respectively), described by a linear Stark formula: $h\Delta\nu = -\Delta\boldsymbol{\mu} \cdot \boldsymbol{E}$, where $\Delta\boldsymbol{\mu}$ is the difference between the permanent electric dipole moments in the excited (e) and ground (g) states,

[1]Department of Physics, Montana State University, Bozeman MT 59717. [2]Department of Chemistry and Biochemistry, Montana State University, Bozeman MT 59717. [3]NEST, Istituto Nanoscienze—CNR and Scuola Normale Superiore, I-56127 Pisa, Italy. [4]Department of Cell Biology and Neuroscience, Montana State University, Bozeman MT 59717. [5]National Insitute of Chemical Physics and Biophysics, Tallinn 12618, Estonia. Correspondence and requests for materials should be addressed to A.R. (email: rebane@physics.montana.edu)
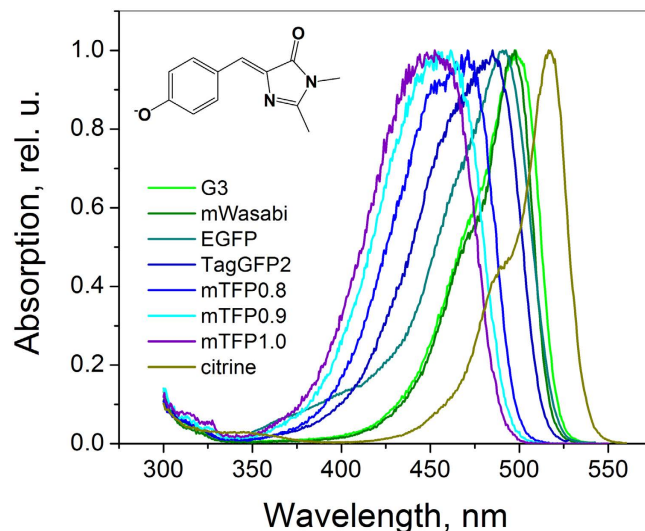
Figure 1. Normalized fluorescence excitation spectra of a series of FPs with the same, GFP-type chromophore in the anionic state.

$\Delta\boldsymbol{\mu} = \boldsymbol{\mu}_e - \boldsymbol{\mu}_g$. The above relation between the field and spectral shift could in principle be used for estimation of the internal field E inside a protein by measuring the shifts of absorption peaks, if $\Delta\boldsymbol{\mu}$ and ν_0 were known. We have previously demonstrated that in a series of red FPs the value of $\Delta\boldsymbol{\mu}$ can be obtained from the two-photon absorption (2PA) cross section, measured at the red side of the S_0-S_1 transition band[14]. Our results have shown however that $\Delta\boldsymbol{\mu}$ significantly changes from one protein to another (for the series with the same chromophore) and therefore the higher order, quadratic Stark effect should be considered. This can be realized by admitting that because of the presence of quite strong local fields inside proteins ($\sim 10^7$ V/cm); $\Delta\boldsymbol{\mu}$ itself becomes a function of E due to the induced (through polarizability) dipole moments. A simple analysis of the quadratic Stark effect in a series of red FPs made it possible to obtain a crude, order-of-magnitude estimation of the fields[14].

However, our initial simplified model based on quadratic Stark shifts did not take into account some important details. First, it considered the chromophore to be one-dimensional, which probably is not very accurate for the red FPs[15]. Second, the field was treated as uniform, i.e. not changing across the π-conjugated system of the chromophore. Although potentially justifiable for the long-range Coulomb interactions, this approximation does not always work for the local short-range interactions (e.g. hydrogen bonding (HB)). Including such interactions is critical, because they can significantly change the values of $\Delta\boldsymbol{\mu}$[16] and transition energy[16–18]. Finally, only pure electronic effects were assumed, i.e. the change of atom positions upon applying the field was disregarded. It is known, however, that structural changes, namely the changes of the bond-length alternation (BLA) inside a chromophore, can result in significant shifts of the optical transition energy and corresponding changes of $\Delta\boldsymbol{\mu}$ values[17–19].

Here we study a set of 26 GFP mutants, all containing the same chromophore in the anionic state, as well as the solution of a model synthetic chromophore, p-hydroxybenzylidene-2,3-dimethylimidazolinone anion (HBDI$^-$) in alkaline D$_2$O. Strictly speaking, mutants/homologs in this series differ for the first chromophore –forming amino acid (Ser65 in the original *av*GFP). However, this amino acid's side chain does not participate to the π-conjugated system, so that the chromophore structure can be considered the same as far as its electronic properties are concerned. We selected the GFP series because their chromophore has simpler, quasi one-dimensional structure, compared to the red FP chromophore. As a first approximation, we consider a simple model, where only pure electronic effects were taken into account (i.e. the positions of atoms are fixed) and the total (short- and long-range) field is considered. To further elucidate the role of the local chromophore environment, we then refined the model by separating the long-range Coulomb interactions from the short-range (hydrogen bonding) specific interactions. This leads us to an expanded "effective" chromophore or, in other words, cluster, which in addition to the chromophore itself involves a number of amino acid residues and water molecules making hydrogen bonds with it. This cluster is embedded in the electrostatic surrounding created by the rest of the protein and water molecules. This latter approach made it possible to evaluate the long-range quasi-uniform fields. Independent evaluations of the total and the long-range fields complement each other and allow us to estimate the local, short-range interactions. Both total and long-range fields measurements are verified by molecular dynamics (MD) simulations.

Results and Discussion

Pure electronic Stark effect model. Total effective electric fields. Here we assume a one-dimensional chromophore, interacting through its electronic system with the electric field of protein

surrounding. We suppose that the chromophore structure, i.e. its bond lengths and angles do not change upon transition from one local surrounding to another. The difference of the chromophore dipole moments $\Delta\boldsymbol{\mu}$ will contain a field-independent (vacuum) part $\Delta\boldsymbol{\mu}_0$ and a field-induced part, $\Delta\boldsymbol{\mu}_{\text{ind}}$. In the point dipole approximation, we can write $\Delta\mu_{\text{ind}} = \Delta\alpha E_{\Delta\mu}$, where $E_{\Delta\mu}$ is the projection of the electric field on the direction of the $\Delta\boldsymbol{\mu}_0$ vector and $\Delta\alpha$ is the change of the chromophore polarizability upon excitation. For simplicity, we will call $E_{\Delta\mu}$ the field in what follows. In our model we assume that $\Delta\alpha$ does not depend on the field (no hyperpolarizability effects). We define a positive molecular x-axis along the direction from the center of the imidazolinone ring to the center of the phenol ring which virtually coincides with the direction of the $\Delta\boldsymbol{\mu}_0$ vector[20–25]. The total change of the permanent dipole moment (projection onto x-axis) then reads[26]:

$$\Delta\mu = \Delta\mu_0 + \Delta\alpha E_{\Delta\mu}. \tag{1}$$

(In our previous paper[14] the coefficient ½ was used in the second term of eq. (1) following an erroneous presentation in some previous literature, see e.g. Ref. 26 A correct consideration[27] should not contain it.) Equation (1) can be solved for the field to obtain:

$$E_{\Delta\mu} = \frac{\Delta\mu - \Delta\mu_0}{\Delta\alpha}. \tag{2}$$

This equation establishes a linear correlation between the experimentally measurable permanent dipole difference and the electric field. Since the field can be non-uniform across the length of the chromophore, we call it effective field, i.e. producing the same effect as the uniform field of magnitude $E_{\Delta\mu}$. This field also shifts the electronic transition frequency relative to that in vacuum via the second-order Stark effect, as follows[27]:

$$h\nu = h\nu_0 - \Delta\mu_0 E_{\Delta\mu} - \frac{1}{2}\Delta\alpha E_{\Delta\mu}^2, \tag{3}$$

where h is the Planck's constant. Substituting the expression for $E_{\Delta\mu}$ (2) into (3) and expressing frequencies in wavenumbers ($\bar{\nu} = \nu/c$, where c is the speed of light), we obtain the relation between the experimentally measurable parameters $\bar{\nu}$ and $\Delta\mu$:

$$\bar{\nu} = \bar{\nu}_0 + \frac{\Delta\mu_0^2}{2hc\Delta\alpha} - \frac{\Delta\mu^2}{2hc\Delta\alpha}. \tag{4}$$

This equation shows that the transition frequency as a function of $\Delta\mu$ is described by a parabola with the extremum found at $\Delta\mu = 0$. Its curvature is defined by $-(2hc\Delta\alpha)^{-1}$.

Figure 2 presents the experimental dependence of the Boltzmann-broadened 0–0 transition frequency on $\Delta\mu$, for the anionic GFP chromophore in a variety of local environments (see Methods for the measurement details). Note that only the 0–0 2PA transition can be safely expected to reflect the dipole change, due to vibronic coupling effect[28].

Fitting of all experimental data to the function $\bar{\nu} = A\Delta\mu^2 + C$ (full black line in Fig. 2) provides the coefficients $A = -(2hc\Delta\alpha)^{-1} = 72 \pm 4 \, \text{cm}^{-1}\text{D}^{-2}$ and $C = \bar{\nu}_0 + \Delta\mu_0^2/2hc\Delta\alpha = 19300 \pm 70 \, \text{cm}^{-1}$. From the first coefficient, we obtain $\Delta\alpha = -35 \pm 2 \, \text{Å}^3$. Substituting this value, as well as the transition frequency measured in the gas phase, $\bar{\nu}_0 = 20725 \, \text{cm}^{-1}$, Ref. 29, into the expression for C, we obtain $\Delta\mu_0 = 4.47 \pm 0.16 \, \text{D}$. Note that both $\Delta\mu_0$ and $\Delta\alpha$ represent some effective values averaged over several different structural variants of the chromophore, e.g. including different bond-length patterns in different classes of proteins or different twisting conformations in water vs. protein environment. Despite this, the experimental values of $\Delta\mu_0$ and $\Delta\alpha$ fall in the range predicted by quantum calculations for the HBDI⁻ chromophore in vacuum (see Table 1 and Supplementary Tables S1 and S2).

By knowing the values of the parameters $\Delta\alpha$ and $\Delta\mu_0$, we find the fields $E_{\Delta\mu}$ from (2) for all the systems studied, see Supplementary Table S3. For some of the proteins, crystallographic structures are available and the corresponding fields can therefore be evaluated theoretically using MD simulations and these results can then be compared to experimental values. Note that the expected outcomes rely only on molecular mechanics (MM) treatment of the protein around the chromophore. In particular, the partial charges in the force field used (see Methods) are obtained by fitting the electrostatic potential (calculated at the Hartree-Fock level) on molecular surfaces and therefore this approach is expected to be sufficiently accurate. For such a comparison we have selected four proteins with very different (in magnitude and direction) electric fields, including YFP-variant citrine (pdb files: 3DPW and 1HUY), EGFP (pdb: 4EUL and 2Y0G), mTFP1.0 (pdb: 2HQK), and mTFP0.7 (pdb: 2OTB).

Figure 3(a) shows the changes of electrostatic potential as a function of distance along the straight line connecting the atoms CA2 and CE2. The potentials at the CE2, CD2, CG2, CB2, and CA2 atoms, closest to the line are shown versus the cumulative distance between the projections of the atoms on that line. Selection of these atoms is justified by the theoretical calculations[25,30,31] demonstrating that the most significant change of electronic density upon the $S_0 \rightarrow S_1$ excitation occurs at the CG2, CB2, CA2, CE2,
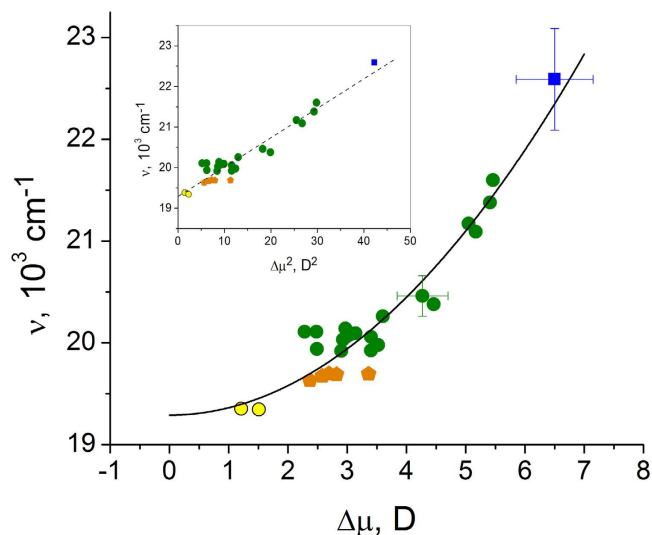
Figure 2. Dependence of the 0–0 transition frequency on the change of the permanent dipole moment for a GFP anionic chromophore in a number of different local environments. The blue square represents the model HBDI⁻ chromophore in alkaline D_2O solution. Standard deviations are shown by bars. Twenty six different FP variants under investigation are sub-divided into three different groups, according to the local structure of their chromophore surrounding: (1) The mutants derived from EGFP and mTFP, where the two chromophore oxygen atoms, i.e. at phenolate and imidazolinone rings, maintain five hydrogen bonds with the surrounding, similar to the original EGFP and mTFP[63–66], are shown by green circles; (2) EGFP-type of mutants containing, among others, the T203I mutation, which causes a reduction of the number of hydrogen bonds from five to four in the chromophore cluster, are shown by orange pentagons; (3) Yellow FP derivatives of GFP containing, among others, the T203Y mutation, which along with the reduction of the number of hydrogen bonds also causes a π-stacking interaction of the chromophore with the Tyr203 phenol[38], are shown by yellow circles. The inset shows the same data in the linearizing, ν versus $\Delta\mu^2$ coordinates.

No.	Chromophore	Environment	$\Delta\mu_x$ D	$\Delta\mu_y$ D	$\Delta\mu_z$ D	$|\Delta\mu|$ D
1	HBDI⁻	vacuum	4.30	0.66	0.11	**4.35**
2	HBDI⁻5 HBs (H_2O)	P: 3 x H_2O I: 2 x H_2O	4.92	1.68	−0.16	**5.20**
3	HBDI⁻ 5 HBs (mTFP)	P: Ser146, His163, H_2O I: Arg95, H_2O	3.04	1.30	−0.01	**3.30**
4	HBDI⁻	vacuum, but chromophore geometry as in 3.	2.86	0.68	−0.01	**2.94**
5	HBDI⁻ 5 HBs (EGFP)	P: Thr203, His148, H_2O I: Arg96, Gln94	2.82	1.10	0.02	**3.02**
6	HBDI⁻ 3 HBs, 4 groups (citrine)	P: His148, H_2O I: Arg96, Gln94	1.55	0.80	−0.25	**1.76**

Table 1. Calculated components and magnitude of $\Delta\mu$ in vacuum and different HB-clusters of the HBDI⁻ chromophore. P: designates the phenolate oxygen local environment; I: designates the imidazolinone oxygen local environment. The entries represent: No. 1 - chromophore in vacuum; No. 2 - cluster (1), as in water solution; No. 3 - cluster (2) as in mTFP; No. 4 - chromophore in vacuum, but with the geometry optimized in cluster (2); No. 5 - cluster (3) as in EGFP; No. 6 - cluster (4) as in citrine. See text for description of cluster structures.

CD2, C2, CE1, CD1, and OH atoms, and that the vector connecting the CA2 and CE2 atoms is virtually co-directed with $\Delta\mu_0$. The three central atoms, CG2, CB2, and CA2, making the bridge between the phenolate and imidazolinone rings are known to be the main players in the charge-transfer process[32–34]. The electric field projected onto the CA2-CE2 direction should therefore be responsible for the induced part of $\Delta\mu$ and for the shifts of transition energy. To obtain $E_{\Delta\mu}$, we took a negative gradient of the calculated electrostatic potential function (i.e. the slopes of the straight lines in Fig. 3(a)). The correlation between the experimentally measured and simulated fields for four proteins is shown in Fig. 3(b). Note
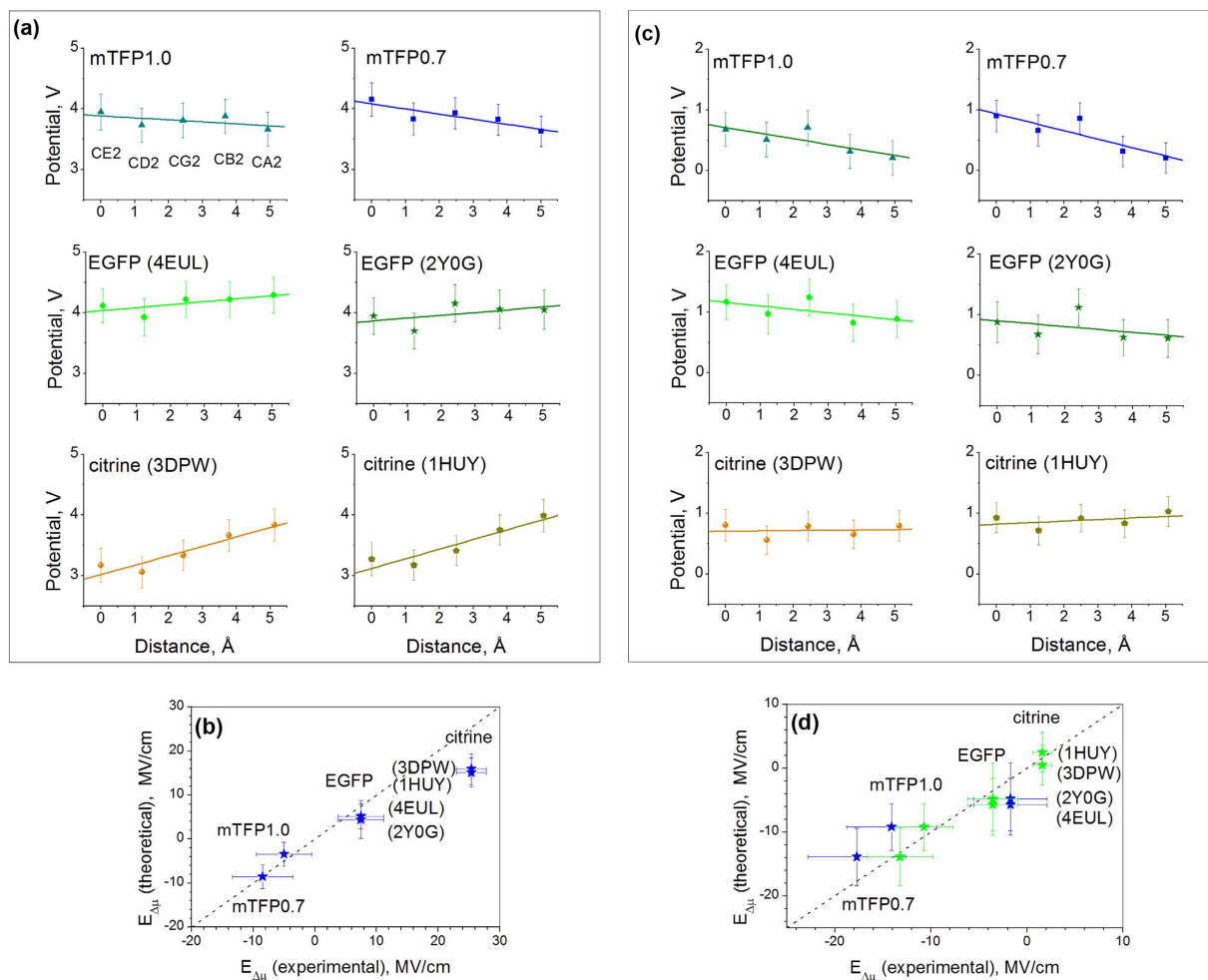
Figure 3. (**a**) MD-simulated electrostatic potentials on the chromophore atoms as a function of the distance along the straight line connecting CE2 and C2 atoms (see Fig. 4 for atom notations) for six selected FP structures. For EGFP and citrine, the two sets of data points correspond to two different pdb files. The linear regressions based on the potentials of the five atoms are shown by straight lines. (**b**) Comparison of experimentally measured electric fields with the fields obtained from the MD simulations of potentials shown in (**a**). Dashed line indicates exact coincidence. (**c**) Same as in (**a**) but when calculating potentials, in addition to the chromophore the contribution from five or four (in case of citrine) amino acids or water molecules in close proximity to the chromophore were also excluded. (**d**) Comparison of electric fields obtained from experimentally measured $\Delta\mu$ values using the parameters $\Delta\mu_{HB}$ and $\Delta\alpha$ either obtained from experiment (blue symbols) or calculated (green symbols) with the fields obtained from the MD simulations of potentials shown in (**c**). Standard deviations are shown by bars in each panel.

that both the magnitude and the sign of the field are reproduced in MD surprisingly well, especially for mTFP0.7, mTFP1.0, and EGFP. The experimental values vary in a wide range, from -8.6 MV/cm in mTFP0.7 to $+7.4$ MV/cm in EGFP, reaching $+25.4$ MV/cm in citrine, which represent typical order of magnitude in proteins[12,13,35–37]. The wide range of the fields observed is due to significant changes of local interactions. In particular, the increase of the component of the field directed from phenolate to imidazolinone (negative x) by ~16 MV/cm when going from EGFP to mTFP0.7 (Fig. 3(b)) can be qualitatively explained by the presence of the unique, positively charged H197 and negatively charged E148 residues near the chromophore in mTFP0.7 [7,18].

Our simple Stark effect model, Fig. 2, also explains the large red shift of yellow FP citrine *vs*. EGFP by the much smaller $\Delta\mu$ value of the former, i.e. 1.5 *vs*. 3.6 D. The intermediate red shifts in T203I mutants (orange pentagons in Fig. 2) are also due to a decrease of $\Delta\mu$. Structurally, in citrine, the key mutation which significantly red-shifts the absorption is T203Y[38]. This substitution reduces the number of hydrogen bonds at the phenolate oxygen from 3 to 2 and, also, creates a new, π-π stacking interaction of the Tyr203 phenol group with the chromophore[38]. It is interesting that the main part of the red shift is due to the loss of a hydrogen bond because even aliphatic substitutions at the same position, i.e. T203V

and T203I produce a similar, although slightly smaller, red shift (see Ref. 39 and Fig. 2). It has been proposed[39] that the extra hydrogen bond, created by the threonine hydroxyl in EGFP, pulls the electron density to the phenolate oxygen, thus stabilizing the ground state and shifting the transition energy up as compared to the EGFP T203I mutant. This interaction is strongly localized near the OH atom and appears to be not fully detectable by averaging the potential changes over the further removed atoms CE2, CD2, CG2, CB2, and CA2, as is done in Fig. 3(a). On the other hand, it significantly changes $\Delta\mu$ of the chromophore and should therefore be observed as a contribution to the effective experimental field. In fact, the difference between the experimental and simulated fields in citrine ~10 MV/cm, see Fig. 3(b), is probably due to this unaccounted short-range interaction. In the next section we refine the model to separate the effects of the short- and long-range fields.

Refined model that includes structural changes of the chromophore. Long-range electric fields. In an attempt to refine the model and elucidate the role of the short-range interactions, which are probably the main cause of discrepancies between the effective experimental and simulated field values (cf. citrine), we now consider a cluster, consisting of the π-conjugated system of the chromophore itself and its closest surrounding. Within the cluster, the chromophore interacts specifically with the inner shell of several chemical groups. In particular, we consider the hydrogen bonding to the OH and O2 atoms of the chromophore as a main source of the short-range interactions. These hydrogen-bonded groups can perturb the initial (vacuum) dipole moments $\boldsymbol{\mu}_g$ and $\boldsymbol{\mu}_e$. As a result, the dipole moment difference of the chromophore in the hydrogen-bonded cluster, $\Delta\mu_{HB}$, can differ from that observed in vacuum. We will further assume that the $\Delta\mu_{HB}$ is invariant for a series of FP mutants with the same number and character of hydrogen bonds in the cluster.

To obtain the total dipole moment difference, $\Delta\mu$, we should include, in addition to $\Delta\mu_{HB}$, a contribution that is induced by the long-range electric field, $E'_{\Delta\mu}$, created by the outer parts of the protein. We use the prime sign to distinguish the long-range part of the field from the total field ($E_{\Delta\mu}$). Using the same approach, as in (1)–(2), and again assuming that $\Delta\alpha$ does not depend on the long- and short-range field variations, we now obtain for the long-range field within a class of proteins with particular $\Delta\mu_{HB}$:

$$E'_{\Delta\mu} = \frac{\Delta\mu - \Delta\mu_{HB}}{\Delta\alpha}.$$

(5)

The unknown parameter $\Delta\mu_{HB}$ can be obtained either from quantum mechanical calculations or from comparison of experimental data with the refined model, describing optical transitions in a series of proteins with the same local surrounding.

Quantum mechanical calculations of $\Delta\mu_{HB}$ and $\Delta\alpha$. For generality, we performed quantum mechanical calculations of the dipole moment differences in vacuum and in four different hydrogen-bonding clusters by using the same method (see Methods for details).

First, the chromophore with five hydrogen-bonded water molecules represented the cluster encountered in water solution (cluster (1)). The cluster (2) included the chromophore with five hydrogen-bonding contacts (5-HB cluster), corresponding to mTFP0.7 structure. It contained His163 (modeled with neutral imidazole group), Ser146 (modeled with the CH_3OH group), and an H_2O molecule, all three bonded to the phenolate oxygen, as well as Arg95 (modeled with positively charged guanidinium group) and another H_2O molecule, both bound to the imidazolinone oxygen. The next cluster (3) corresponded to the EGFP where His148, Thr203, and a water molecule were H-bonded to the phenolate oxygen and Arg96 and Gln94 were H-bonded to the imidazolinone oxygen. The cluster (4), corresponding to citrine, contained the chromophore with His148 and H_2O both H-bonded to the phenolate oxygen and R96 (H-bonded to imidazolinone oxygen) and Q94 (probably weakly or not H-bonded, but very close to imidazolinone oxygen). The Q94 is retained in the citrine cluster for easier comparison with EGFP.

We select the molecular frame, such that the x-axis is directed from the geometrical center of the imidazolinone ring to the geometrical center of the phenol ring that is virtually parallel to the (calculated) transition dipole vector. The y-axis is directed orthogonally and lies in the average molecular plane, as shown in the Fig. 4. The z-axis makes a right triad with x and y. The results of the calculations are presented in Table 1.

The calculation shows that the $\Delta\boldsymbol{\mu}_{HB}$ vector is directed almost perfectly along the x-axis, similarly to what was found in previous calculations for the HBDI⁻ chromophore in vacuum, water, and GFP protein (Supplementary Table S1), thus justifying the one-dimensional model of the chromophore.

Note that the calculated value of $\Delta\mu$ in vacuum (4.35 D) is in good agreement with the experiment (4.47 ± 0.16 D). The addition of five hydrogen-bonded water molecules to the chromophore slightly increases the $\Delta\mu$ magnitude (~5.2 D). On the other hand, the $\Delta\mu$ value decreases (compared to vacuum) in the mTFP-derived 5-HB cluster to 3.3 D (entry 3) which, according to our analysis, can be mainly attributed to the positively-charged Arg95. Interestingly, removing all the surrounding groups, but keeping the chromophore structure the same as it was found after optimization in the cluster, produces only minor change of the $\Delta\mu$, i.e. from 3.3 to 2.9 D, cf. entries 3 and 4 of Table 1. This important result demonstrates that in this case the change of $\Delta\mu$ in cluster *vs.* vacuum is primarily due to the change of the chromophore structure, namely to a different bond-length alternation pattern (cf. Ref. 17), and
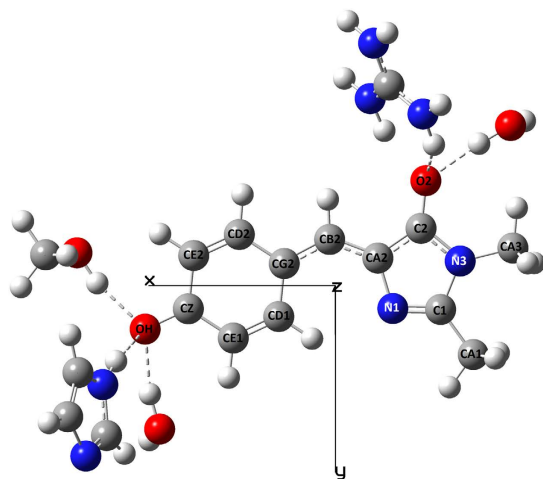
Figure 4. Schematic of the cluster, using mTFP0.7 structure (pdb file 2OTB) as an example that contains the HBDI⁻ chromophore and five hydrogen-bonded groups.

that the electronic redistribution is coupled to the structural change. Another interesting result is that in the citrine cluster, where there are only three hydrogen bonds (entry No. 6), the $\Delta\mu_{HB}$ value is further significantly reduced compared to that observed in the 5-HB clusters.

To obtain the $\Delta\alpha$ value, the dependence of $\Delta\mu$ on $E_{\Delta\mu}$ was calculated by applying a uniform electric field along the x-axis of the chromophore (Methods and Supplementary Fig. S1). The resulting value, $\Delta\alpha = -49\,\text{Å}^3$, is in reasonable agreement with the experimental one ($-35\,\text{Å}^3$). Using the calculated parameters of $\Delta\alpha$ and $\Delta\mu_{HB}$ (Table 1), as well as experimental values of $\Delta\mu$, we evaluated the long-range field $E'_{\Delta\mu}$ by using eq. (5). For four selected proteins the results are presented in Fig. 3d (abscissa, green stars) and for the rest of the proteins – in Supplementary Table S3.

Experimental evaluation of $\Delta\mu_{HB}$. We also carried out an independent evaluation of the parameter $\Delta\mu_{HB}$ for the subset of 5-HB proteins using experimental measurements. To this end, we considered the dependence of transition frequency on the electric field by separating the effects of structural and pure electronic responses. First, we assume that both short- and long-range electric fields perturb molecular structure such that the alternation of the single and double bond lengths changes accordingly. To obtain the dependence of transition frequency on the field and hence on $\Delta\mu$, the resonance theory was applied in the simple two-forms two-states (2F2S) implementation[40–42] adapted to the GFP chromophore in[17–19,28,43]. This model predicts the quadratic dependence of the optical transition frequency on the value of $\Delta\mu$, Ref. 19. We then include the pure electronic interaction of the chromophore system with the long-range electrostatic field through the second-order Stark effect. As a result, the following dependence of the transition frequency on $\Delta\mu$ was obtained (see Supplementary Discussion):

$$\bar{\nu} = A\Delta\mu^2 + B\Delta\mu + C, \tag{6}$$

with

$$A = \frac{1}{hc}\left(\delta - \frac{1}{2\Delta\alpha}\right), \tag{7}$$

$$B = -\frac{2\delta\Delta\mu_c}{hc}, \tag{8}$$

$$C = \left(\bar{\nu}_I + \frac{\Delta\mu_{HB}^2}{2hc\Delta\alpha} + \frac{\delta\Delta\mu_c^2}{hc}\right), \tag{9}$$

where $\bar{\nu}_I$ is the parameter describing the coupling between the two resonating forms[19], $\delta = hc\bar{\nu}_I/8\eta^2\mu^2$, η is a dimensionless parameter whose physical meaning consists in the "screening" of the charge-transfer transition via redistribution of those electrons clouds which are not directly involved in the electron transfer related to $\Delta\mu$[19,44], μ is the transition dipole moment between the ground and excited states, $\Delta\mu_c$ is a constant, corresponding to the difference between the dipole moments in the excited and ground states in a particular chromophore geometry corresponding to the effective bond length alternation equal

Parameter	$\overline{\nu}_I$ cm^{-1}	η	μ, D	δ, Å^{-3}	$\Delta\mu_c$, D	$\Delta\mu_{HB}$, D
Experiment	20,620	0.57	6.9	0.033	3.69 ± 0.14	3.4 ± 0.2
Calculation (Ref.)	20,462 (20)	0.49 (20)	10.6 (this work)			3.0–3.3 (this work)

Table 2. Experimental and calculated model parameters, describing the transition frequency as a function of $\Delta\mu$ in a series of FPs where the effective chromophore is presented as a cluster containing the HBDI$^-$ molecule and five hydrogen-bonded groups, i.e. representing mTFP- and EGFP-derived mutants (see Fig. 4). Definition of parameters is presented in the text. Standard deviation of $\Delta\mu_c$ and $\Delta\mu_{HB}$ are shown in the last two columns.

to 0 (see Supplementary Fig. S2). The three coefficients, A, B, and C were found from the best fit of the $\overline{\nu}$ vs $\Delta\mu$ dependence for a subset of proteins with 5 HBs in the cluster to a second order polynomial (see Supplementary Fig. S3). Since there are four unknown parameters ($\overline{\nu}_I$, δ, $\Delta\mu_c$, and $\Delta\Delta\mu_{HB}$) in the system of three equations (7)–(9), we use an additional equation which is obtained by considering the 2F2S model in the zero field (vacuum), i.e. $E'_{\Delta\mu} = 0$ and $\Delta\mu = \Delta\mu_0$. This equation reads (see Supplementary Discussion):

$$\overline{\nu}_I + \frac{\delta\left(\Delta\mu_0 - \Delta\mu_c\right)^2}{hc} = \overline{\nu}_0, \tag{10}$$

where $\overline{\nu}_0 = 20725$ cm^{-1} is the transition frequency in vacuum[29]. The resulting experimental parameters involved in the model were obtained by solving eqs. (7)–(10) (See Supplementary Discussion) and are presented in Table 2. Most of them were also independently calculated quantum-mechanically (Table 2) and agree well with the experiment.

Combining the values of $\Delta\mu_{HB} = 3.4$ D and $\Delta\alpha = -35$ Å^3, both found experimentally, the fields $E'_{\Delta\mu}$ were calculated according to (5) (Supplementary Table S3).

Comparison of the long-range fields, obtained from $\Delta\mu$ and from MD simulations. To check these two sets of data (with $\Delta\mu_{HB}$ and $\Delta\alpha$ either calculated or found from experimental measurements) against the MD simulations, the same group of mutants was selected as in Fig. 3(a)–(b). Now the charges on the chromophore and 5 amino acids/water molecules (4 in case of citrine) around it were disregarded in calculating the potentials. Figure 3(c) shows the variation of potentials along the straight line connecting the CE2-CD2-CG2-CB2-CA2 chain of atoms, which provided the fields for all the proteins.

Figure 3(d) compares the MD simulated fields (*y*-axis) with the fields obtained from model eq. (5). The correlation between the MD simulations and the model is good (particularly for the calculated $\Delta\mu_{HB}$ and $\Delta\alpha$, green stars). For citrine it is improved significantly, as compared to the case of the total effective field, cf. Figure 3(b). The long-range electric fields inside the proteins under study vary in the range from ~1.6 MV/cm in citrine to −13 MV/cm in mTFP 0.7. This variation is smaller than that of the total effective field, Fig. 3(b), which emphasizes the importance of the short-range interactions in the wide spectral tuning of the FP. The direction of the long-range field, found in the four mutants under study is opposite to that of $\Delta\mu$, i.e. $E'_{\Delta\mu}$ is directed from the phenolate to the imidazolinone. Note that in the 5-HB proteins, mTFP0.7, mTFP1.0, and EGFP, the field was shifted to more negative values by ~9 MV/cm, when the cluster was discarded (cf. Figure 3 (b,d), blue stars). This is because the positively charged R96 residue (EGFP notation), being a part of the cluster in the three proteins, created the short-range field in the positive direction, i.e. from imidazolinone to phenolate, see Fig. 4.

In the case of citrine, the long-range field value is much better reproduced by MD simulations and it is also much closer to that of EGFP, as compared to what was observed for the total effective field, cf. Figure 3(d,b). This reflects the fact that the large part of structural changes between the two mutants occur at a short range. In particular, the reduced number of hydrogen bonds in citrine vs. EGFP results in a significant decrease of $\Delta\mu$, from 3.6 D (EGFP) to 1.5 D (citrine), which can be simulated by a large increase of the short-range contribution to the effective field. Note that the much lower value of $\Delta\mu$ in citrine is consistent with the previous observation of the smaller static first hyperbolarizability in EYFP compared to EGFP[45,46]. A little upward shift of the long-range field, by ~5 MV/cm, in citrine *vs.* EGFP, cf. green stars in Fig. 3(d), can be tentatively attributed to the dipole-induced contribution of the Tyr203 phenol group.

Field-induced broadening of the absorption spectrum. Simple inspection of Fig. 1 suggests that the blue shift of absorption in the series of mutants is accompanied by systematic broadening of the spectrum. Since the more blue-shifted spectra correspond to larger $\Delta\mu$ (Fig. 2), we hypothesized that the broadening is due to Stark-induced variations of transition frequencies in response to stochastic fluctuations of local fields. If we assume that the field's fluctuations obey the normal distribution with the
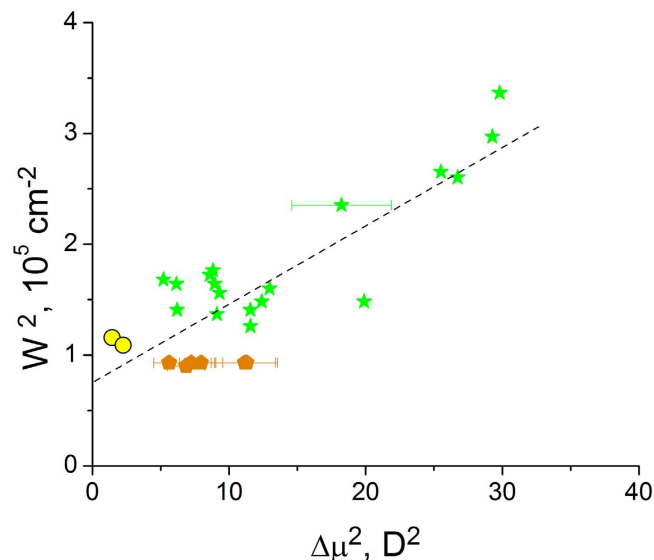
Figure 5. Dependence of the square of the spectral width (standard deviation of the Gaussian envelope, see Methods) of the 0–0 1PA transition on the square of the change of permanent dipole moment for a series of 26 GFPs. The colors of symbols are the same as in Fig. 2.

standard deviation $\delta E_{\Delta\mu}$ and that $\delta E_{\Delta\mu}$ does not vary much from one protein to another (which is supported by MD simulations, see Supplementary Tables S4 and S5), then larger $\Delta\mu$ should produce a larger spread of transition frequencies. In this case, the field-dependent part of the spectral broadening, $\delta\bar{\nu}$, can be obtained by using Eqs. (3) and (2),

$$\delta\bar{\nu} = \frac{\Delta\mu}{hc}\delta E_{\Delta\mu}. \tag{11}$$

If, in addition to $\delta\bar{\nu}$, there is also a field-independent broadening, due to e.g. low-frequency vibrations, which is described by a Gaussian envelope with standard deviation γ, then the total linewidth of the 0–0 transition will read:

$$w = \left[\gamma^2 + \left(\frac{\delta E_{\Delta\mu}}{hc}\right)^2\Delta\mu^2\right]^{1/2}. \tag{12}$$

Figure 5 shows an experimental correlation of w^2 and $\Delta\mu^2$. From the slope and intercept of the linear regression, we obtain $\delta E_{\Delta\mu} = (5 \pm 0.3)\,MV/cm$ and $\gamma = 270 \pm 25\,cm^{-1}$, respectively. The first value compares well with the standard deviation of the fields fluctuations found in MD simulations ($\sim$6–7 MV/cm, see Supplementary Tables S4 and S5). The second value is close to the linewidth of the 0–0 transition of the bare chromophore in vacuum ($310 \pm 45\,cm^{-1}$)[29]. This justifies our assumption that the protein-to-protein variations of spectral widths can be explained by the electric field-induced variations of the Stark shifts of individual transitions.

Specific case: HBDI⁻ chromophore in water. Solvent reaction field E = E_R. The 0–0 transition of the HBDI⁻ in D_2O solution is significantly shifted to higher frequencies compared to all the proteins and the corresponding $\Delta\mu$ value reaches the largest value, 6.5 D, Fig. 2. Several high-level quantum calculations predict similar $\Delta\mu$ values for HBDI⁻ in water (Supplementary Table S1). The CASPT2/CASSCF method treating solvent molecules explicitly[16,22] provides a very good agreement with the experiment. Simulation of the solvent with a polarized continuum model gives a slightly overestimated result, i.e. 8.6 D[24]. Our own quantum mechanical calculations show that the addition of five hydrogen-bonded water molecules to a bare chromophore (three at the phenolate oxygen and two at the imidazolinone oxygen, cf. Ref. 22) slightly increases $\Delta\mu$, from $\simeq$4.4 D (in vacuum, both experimental and theoretical) to 5.2 D. Therefore the increase of $\Delta\mu$ from 4.4 D in vacuum (from 5.2 D in cluster) to $\Delta\mu = 6.5\,D$ in bulk water may be attributed to the effect of a polarizable water solution around the chromophore (or cluster). The corresponding part of the dipole moment change, $\Delta\mu_{ind} = 2.1\,D$ (1.3 D), induced by the reaction field E_R of the polarized solvent, can be presented as follows, cf. (1),

$$\Delta\mu_{ind} = \Delta\alpha E_R \cdot \tag{13}$$

Here both $\Delta\mu_{ind}$ and E_R represent the projections of the corresponding vectors on the direction of $\Delta\mu_0$ ($\Delta\mu_{HB}$). Using experimental values of $\Delta\alpha = -35\,\text{Å}^3$ and $\Delta\mu_{ind} = 2.1\,\text{D}$ (1.3 D), we find $E_R = -18\,\text{MV/cm}$ ($-11\,\text{MV/cm}$).

For comparison, we also use the classical polarizable dielectric model to estimate the reaction field of the solvent. In this model, the field E_R is created by a chromophore with the ground-state dipole moment μ_g and polarizability α_g, that is placed in the center of a spherical cavity with radius a possessing the internal dielectric constant ε_{in}, which in turn is immersed in a solvent with the dielectric constant ε. The field is assumed to be the same for the ground and excited states of a molecule and is presented as[47]:

$$E_R = \frac{\mu_g f}{a^3 - \alpha_g f}, \tag{14}$$

where f is the cavity field factor, described as[47,48]:

$$f \cong \frac{2}{\varepsilon_{in}} \frac{(\varepsilon - \varepsilon_{in})}{2\varepsilon + \varepsilon_{in}} \cdot \tag{15}$$

An estimation, using $\varepsilon = 79$ and $\varepsilon_{in} = 2$[47,48], as well as experimentally determined parameters (see Supplementary Discussion): $a = 4.7\,\text{Å}$ (for the cluster of the HBDI$^-$ chromophore with 6 hydrogen-bonded water molecules), $\alpha_g = 22\,\text{Å}^3$, and literature value $\mu_g = -11 \pm 2\,\text{D}$[22,31,49], yields $E_R = -17\,\text{MV/cm}$. The latter value agrees well with those obtained above from the field-induced dipole moment change in cluster ($-11\,\text{MV/cm}$) and in vacuum ($-18\,\text{MV/cm}$). Due to symmetry reasons, the field $\boldsymbol{E}_R$ is directed along the μ_g vector. Since in HBDI$^-$ chromophore the permanent dipole moment decreases upon excitation[20,50], the $\Delta\boldsymbol{\mu}$ is antiparallel to $\boldsymbol{\mu}_g$, and therefore $\boldsymbol{E}_R$ is antiparallel to $\Delta\boldsymbol{\mu}$. This means that $\boldsymbol{E}_R$ found for water has the same direction (in the chromophore frame) and similar amplitude as the long-range field $\boldsymbol{E'}_{\Delta\mu}$ inside the mTFP0.7 protein.

In conclusion, we have evaluated local electric fields in a series of fluorescent proteins with the same anionic GFP chromophore by using simultaneous measurements of 1PA and 2PA cross sections and the 0–0 transition frequency. In the first approximation, we considered the chromophore interacting with the rest of the protein and obtained the total effective protein field. In the second, - we considered a subset of FPs with similar local surrounding and selected a hydrogen-bonded cluster interacting with further-separated layers of the protein, which made it possible to obtain the long-range part of the field. Independent evaluations of the total and the long-range fields complement each other and provide the local, short-range field. We have shown that both total and long-range fields are well reproduced by MD simulations, which justifies our method of measurement. Our model quantitatively explains the spectral shifts of yellow and teal FPs versus EGFP as well as the spectral position of the isolated chromophore in water solution.

Methods

Measurements of $\Delta\mu$.   The method of the measurement of $\Delta\mu$ using two-photon absorption (2PA) spectroscopy was described previously[14,51]. Briefly, first the one-photon absorption spectrum (obtained in the forms of fluorescence excitation spectrum) of a protein was fitted with four or five Gaussians of the same width (three Gaussian fit was used for the HBDI- chromophore in D$_2$O), see Supplementary Fig. S4. The width was systematically varied to obtain the minimum χ^2. This resulted in the best fit of the 1PA spectrum, yielding three important parameters: spectral position of the lowest-energy transition ν_{0-0}, its corresponding peak extinction coefficient $\varepsilon(0-0)$, and its spectral width (Gaussian standard deviation) w. As a next step, the two-photon absorption spectrum presented as the σ_2 value versus transition frequency (i.e. twice the laser frequency) was fitted to the same number of Gaussian peaks where the positions and amplitudes of the peaks were free, but the widths were fixed and equal to w (i.e. found from fitting of the 1PA spectrum). The two-photon cross section $\sigma_2(0-0)$, corresponding to the lowest-energy Gaussian peak was obtained from the best fit. The $\Delta\mu$ value was then calculated as follows[14]:

$$|\Delta\mu| = \left(\frac{5}{4(1 + 2\cos^2\gamma)} \frac{hc^2 N_A}{\pi 10^3 \ln 10} \frac{n}{f_{opt}^2} \frac{\bar{\nu}_{0-0}}{\varepsilon(0-0)} \sigma_2(0-0) \right)^{1/2}, \tag{16}$$

where γ is the angle between $\boldsymbol{\mu}$ and $\Delta\boldsymbol{\mu}$ vectors, which in the case of GFP chromophore is $\approx 20^0$ (see Ref. 52), $f_{opt} = (n^2 + 2)/3 = 1.26$ is the Lorentz local field factor calculated for water.

The two-level approximation used in the derivation of eq. (16) was justified by quantum chemical calculations of the sum-over-states expression of the 2PA cross section for HBDI- chromophore. For 30 different arrangements (including HBDI$^-$ chromophore in vacuum and in different clusters representing local HB-surrounding of EGFP, TFP, and citrine mutants), ZINDO calculations of $\Delta\mu$ and 2PA cross

sections show that the contribution of intermediate states higher than S_1 to the σ_2 value was not larger than 5% (average ~ 4%). A few contributions of 10–20% were found only when the 2PA was weak due to a small dipole change ($\Delta\mu \leq 1\,D$).

The 1PA and 2PA spectra of all of the FPs studied here were published elsewhere[53,54]. The 2PA spectrum of the HBDI$^-$ chromophore (Supplementary Fig. S4) was measured using the femtosecond nonlinear transmission (NLT) setup described in[55]. Fluorescein in water (pH11) was used for the 2PA standard (21.3 mM in 0.5 cm cuvette). The sample solution was a 10 cm path length of 19.6 mM HBDI$^-$ in $D_2O + 0.1\,M$ NaOD. The Rayleigh length of the beam for the measurement was > 50 cm for all wavelengths. The maximum two-photon cross section obtained for this sample (40 GM) is close to what was observed previously for a similar molecule[56].

Quantum calculations.   Beginning with the X-ray structure 2OTB.pdb for the mTFP 0.7 conformation, the chromophore and the surrounding residues His163, Ser 146, Wat 368, Arg 95, and Wat 361 were read into GaussView 5.08, with no alteration of coordinates except addition of H atoms. The resulting structures were truncated and terminated with methyl groups, yielding the structure in Fig. 4. The same approach was used for EGFP (4EUL.pdb) and citrine (3DPW.pdb) clusters. The geometry of the chromophore and any water molecules was optimized using m062x/6–31 g(d), while the non-water members of the cluster were frozen at the original pdb coordinates. CHARMM27[57] point charges were placed at the resulting cluster atom coordinates, and $\Delta\mu$ values, along with transition energies and transition moments, were obtained for the lowest 30 excited states from Zerner's spectroscopically calibrated INDO/S2-CIS (ZINDO)[58,59] with oxygen parameters from the Truhlar group[60]. 196 excited singlet configurations were included using Mataga-Nishimoto electron repulsion screening and the original CNDO/S overlap factors. The cluster was represented with point charges because we found that including the entire cluster in the ZINDO calculations introduced small amounts of contamination from low-lying charge transfer states involving the nearby positively charged Arg into the lowest excited state, S_1. The large $\Delta\mu$ values of the CT states made the computed $S_0 \rightarrow S_1$ $\Delta\mu$ values erratic and unreliable.

Optimizing the chromophore geometry in vacuum or in the cluster showed the resulting $\Delta\mu$ to be reasonably insensitive to method and basis set, see Supplementary Tables S6 and S7.

To obtain the $\Delta\alpha$ value, the dependence of $\Delta\mu$ on $E_{\Delta\mu}$ was calculated by applying a uniform electric field to the chromophore (Supplementary Fig. S1), using a combination of DFT and ZINDO. The geometry of the chromophore was optimized in the ground state using m062x/6–31 g(d) at each value of the field. The DFT+field-optimized geometries were then used in ZINDO calculations of $\Delta\mu$ vs. field between −15 and 0 MV/cm, i.e. corresponding to the range of the actual fields. The slope of the resulting linear regression describing the $\Delta\mu$ ($E_{\Delta\mu}$) yields $\Delta\alpha = -49\,\text{Å}^3$, i.e. similar to experiment ($-35\,\text{Å}^3$). Note that the absolute $\Delta\alpha$ values obtained from DFT methods using different functionals are 1.6–4 times smaller than the experimental one (Supplementary Table S2).

Starting from the corresponding PDB structure (citrine: 1HUY[61], 3DPW[62]; EGFP: 2Y0G[63], 4EUL[64]; mTFP0.7: 2OTB[65]; mTFP1: 2HQK[66]), the first chain was selected in case of dimer or tetramer complexes, preserving chrystallographic water molecules within 4 Å of the protein chain and hydrogen atoms were added by the pdb2gmx utility of Gromacs. Regarding the protonation state of some relevant residues, in citrine and EGFP the protonation state of Glu222 was set neutral, according to the accepted model, whereas all other Glu were considered anionic. All Histidine residues were set neutral, except His197 of mTFP0.7 and mTFP1, which was set cationic (protonated). In EGFP and citrine His25, 148, 169, 181, 199, and 217 were set as δ (i.e. with the proton in their δ nitrogen atom), and His77, 81, and 139 as ε (i.e. with the proton in their ε nitrogen atom). In mTFP0.7 and mTFP1 His25, 123, 163 and 204 (172, 173, 193) were set as δ (ε). The proteins were solvated in a truncated octahedron box of initial side length 75 Å, the complete systems containing, beside the protein, ~9500–9600 total water molecules, 20 Cl- ions and the following number of Na+ ions: 27 in 4EUL, 2Y0G and 1HUY, 23 in 2OTB, 22 in 2HQK, to neutralize the system. The different numbers of Na+ in the two citrine models 1HUY and 3DPW are due to the presence of the solvent exposed Q80R mutation in the former structure.

The amber99SB*-ILDN force field was used for the simulations[67,68]. The force field of the anionic chromophore was adapted from Ref. 69. Periodic boundary conditions were applied and electrostatics was treated using the smooth particle mesh Ewald method with a grid spacing of 1.2 Å and a 10 Å real-space cutoff. The simulation protocol consisted in initial 400 steps conjugate-gradient minimization with restraints on the protein non-hydrogen atoms (spring force 5 kJ mol^{-1} Å^{-2}), followed by three 200-ps runs during which the restraints spring force constant was set at 5, 3 and 1 kJ mol^{-1} Å^{-2}. In the following 2-ns production runs maintaining 1 kJ mol^{-1} Å^{-2} restraints, the system configurations were saved each ps. In all MD runs a 2 fs time step was used, LINCS was applied to constrains covalent bond lengths involving hydrogen atoms[70], and the neighbor list is updated every 10 steps. Constant 300 K temperature and 1 kbar pressure were maintained by v-rescale thermostat[71] (with a coupling of $\tau_T = 1$ ps) and Parrinello-Rahman barostat ($\tau_P = 4$ ps). The simulations were performed with the Gromacs 4.6.3 package[72].

The electrostatic potential V at site **R** was calculated for each snapshot (disregarding the first 500 ps of simulation) using a homemade script as the sum of the contribution of each atom's Amber partial charge q_i, i.e.

11

$$V\left(\overrightarrow{R}\right) = \sum_{i'} \frac{q_i}{\left|\overrightarrow{R} - \overrightarrow{r}_i\right|} \tag{17}$$

r_i being the position of atom i. The sum did not include either (a) the bare chromophore (atoms as in HBDI⁻ model, including up to the carbon atoms CA1 and CA3 of the methyl groups connected to the imidazolidinone ring), or (b) a cluster containing HBDI⁻ plus in citrine Gln94, Arg96, His148 and phenolate H-bonded water molecule; in EGFP same residues as in citrine plus Thr203; in the mTFPs Arg95,Ser146,His163, two water molecules one H-bonded to the phenolate, the other to the carbonyl of the imidazolidinone. Before calculating V the protein was centered in the box, and periodic boundary conditions were not accounted for, with the assumption that spurious surface effects are averaged out by the free-moving water molecules and ions. Comparison with the full Ewald sum calculation (using the PMEPot routine of VMD with an Ewald factor of 1 Å and a 0.2 Å grid[73]) was performed on one snapshot yielding relative errors smaller than 5%. The potentials at the chromophore atom positions as generated by all protein and water molecules (crystallographic and added to the simulation box) were averaged over the full simulation time (1.5 ns). Fluctuations (calculated as the standard deviation) of the electrostatic potential during MD were around 0.25 V, with a maximum of 0.38 V (on atom O2 of EGFP 4EUL).

References

1. Warshel, A. *et al.* Electrostatic basis for enzyme calalysis. *Chem. Rev.* **106,** 3210–3235 (2006).
2. Fried, S. D., Bagchi, S. & Boxer, S. G. Extreme electric fields power catalysis in the active site of ketosteroid isomerase. *Science* **346,** 1510–1514 (2014).
3. Asamoah, O. K., Wuskell, J. P., Loew, L. M. & Bezanilla, F. A fluoremetric approach to local electric field measurements in a voltage-gated ion channel. *Neuron* **37,** 85–97 (2003).
4. Starace, D. M. & Bezanilla, F. A proton pore in a potassium channel voltage sensor reveals a focused electric field. *Nature* **427,** 548–553 (2004).
5. Fujimoto, K., Hasegawa, J.-Y. & Nakatsuji, H. Origin of color tuning in human red, green, and blue cone pigments: SAC-CI and QM/MM study. *Chem Phys. Lett.* **462,** 318–320 (2008).
6. Tsien, R. Y. The green fluorescent protein. *Annu. Rev. Biochem.* **67,** 509–544 (1998).
7. Ai, H.-W., Olenych, S. G., Wong, P., Davidson, M. W. & Campbell, R. E. Hue-shifted monomeric variants of Clavularia cyan fluorescent protein: identification of the molecular determinants of color and applications in fluorescence imaging. *BMC Biol.* **6,** 13 (2008).
8. Topol, I., Collins, J. & Nemukhin A. Modeling spectral tuning in monomeric teal fluorescent protein mTFP1. *Biophys. Chem.* **149,** 78–82 (2010).
9. List, N., Olsen, J. M. H., Jensen, H. J. A., Steindal, A. H. & Kongsted, J. Molecular-level insight into the spectral tuning mechanism of the DsRed chromophore. *J. Phys. Chem. Lett.* **3,** 3513–3521 (2012).
10. Kaila, V. R. I., Send, R. & Sundholm, S. Electrostatic spectral tuning mechanism of the green fluorescent protein. *PCCP* **15,** 4491–4495 (2013).
11. Beerepoot, M. T. P. *et al.* A polarizable embedding DFT study of one-photon absorption in fluorescent proteins. *PCCP* **15,** 4735–4743 (2013).
12. Bublitz, G. U. & Boxer, S. G. Stark spectroscopy: Applications in chemistry, biology, and materials science. *Annu. Rev. Chem. Phys.* **48,** 213–242 (1997).
13. Boxer, S. G. Stark realities. *J. Phys. Chem. B* **113,** 2972–2983 (2009).
14. Drobizhev, M., Tillo, S., Makarov, N. S, Hughes, T. E. & Rebane, A. Color hues in red fluorescent proteins are due to internal quadratic Stark effect. *J. Phys. Chem. B* **113,** 12860–12864 (2009).
15. Yan, W., Zhang, L. & Xie, D. Electronic excitations of green fluorescent protein: Modeling solvatochromic shifts of red fluorescent protein chromophore model compound in aqueous solutions. *J. Phys. Chem. B* **111,** 14055–14063 (2007).
16. Sinicropi, A., Andruniow, T., Ferré, N., Basosi, R. & Olivucci, M. Properties of the emitting state of the green fluorescent protein resolved at the CASPT2//CASSCF/CHARMM level. *J. Am. Chem. Soc.* **127,** 11534–11535 (2005).
17. Laino, T., Nifosì, R. & Tozzini, V. Relationship between structure and optical properties in green fluorescent proteins: A quantum mechanical study of the chromophore environment. *Chem. Phys.* **298,** 17–28 (2004).
18. Amat, P. & Nifosì, R. Spectral "fine" tuning in fluorescent proteins: The case of the GFP-like chromophore in the anionic protonation state. *J. Chem. Theory Comput.* **9,** 497–508 (2013).
19. Olsen, S. & McKenzie, R. H. Bond alternation, polarizability, and resonance detuning in methine dyes. *J. Chem. Phys.* **134,** 114520 (2011).
20. Martin, M. E., Negri, F., Ollivuci M. Origin, nature, and fate of the fluorescent state of the green fluorescent protein chromophore at the CASPT2//CASSCF resolution. *J. Am. Chem. Soc.* **126,** 5452–5464 (2004).
21. Filippi, C., Buda, F., Guidoni, L. & Sinicropi, A. Bathochromic shift in green fluorescent protein: A puzzle for QM/MM approaches. *J. Chem. Theor. Comput.* **8,** 112–124 (2012).
22. Zhang, L., Xie, D. & Zeng, J., Electronic excitations of green fluorescent proteins: Modeling solvatochromic shifts of chromophore model compounds in solutions. *J. Theor. Comput. Chem.* **5,** 375–390 (2006).
23. Wanko, M., García-Risueño, P. & Rubio, A. Excited states of the green fluorescent protein chromophore: Performance of ab initio and semiempirical methods. *Phys. Stat. Sol. B* **249,** 392–400 (2012).
24. Altoe, P., Bernardi, F., Garavelli, M., Orlandi, G. & Negri, F. Solvent effects on the vibrational activity and photodynamics of the green fluorescent protein chromophore: A quantum-chemical study. *J. Am. Chem. Soc.* **127,** 3952–3963 (2005).
25. Nifosì, R., Amat, P. & Tozzini V. Variation of spectral, structural, and vibrational properties within the intrinsically fluorescent proteins family: A density functional study. *J. Comput. Chem.* **28,** 2366–2377 (2007).
26. Berlin, Y., Burin, A., Friedrich, J. & Kohler, J. Spectroscopy of proteins at low temperature. Part I: Experiments with molecular ensembles. *Phys. Life Sci.* **3,** 262–292 (2006).
27. Atkins, P. W. & Friedman R. S. *Molecular Quantum Mechanics.* (Oxford Univ. Press, Oxford, UK, 2010).
28. Drobizhev, M., Makarov, N. S., Tillo, S. E., Hughes, T. E. & Rebane A. Describing two-photon absorptivity of fluorescent proteins with a new vibronic coupling mechanism. *J. Phys. Chem. B* **116,** 1736–1744 (2012).
29. Forbes, M. W. & Jockusch, R. A. Deactivation pathways of an isolated green fluorescent protein model chromophore studied by electronic action spectroscopy. *J. Am. Chem. Soc.* **131,** 17038–17039 (2009).

30. Helms, V., Winstead, C. & Langhoff, P. W. Low-lying electronic excitations of the green fluorescent protein chromophore. *J. Mol. Struct. (Theochem)* **506,** 179–189 (2000).

31. Wan, S. *et al.* Photoabsorption of green and red fluorescent protein chromophore anions *in vacuo. Biophys. Chem.* **129,** 218–223 (2007).

32. Olsen, S. & Smith, S. C. Bond selection in the photoisomerization reaction of anionic green fluorescent protein and kindling fluorescent protein chromophore models. *J. Am. Chem. Soc.* **130,** 8677–8689 (2008).

33. Epifanovsky, E., Polyakov, I., Grigorenko, B., Nemukhin, A. & Krylov A. I. Quantum chemical benchmark studies of the electronic properties of the green fluorescent protein chromophore. 1. Electronically excited and ionized states of the anionic chromophore in the gas phase. *J. Chem. Theor. Comput.* **5,** 1895–1906 (2009).

34. Epifanovsky, E., Polyakov, I., Grigorenko, B., Nemukhin, A. & Krylov A. I. The effect of oxidation on the electronic structure of the green fluorescent protein chromophore. *J. Chem. Phys.* **132,** 115104 (2010).

35. Geissinger, P., Kohler, B. E. & Woehl, J. C. Electric-field and structure in the myoglobin heme pocket. *J. Phys. Chem.* **99,** 16527–16529 (1995).

36. Callis, P. R. & Burgess B. K. Tryptophan fluorescence shifts in proteins from hybrid simulations: An electrostatic approach. *J. Phys. Chem. B* **101,** 9429–9432 (1997).

37. Manas, E. S., Wright, W. W., Sharp, K. A., Friedrich, J. & Vanderkooi, J. M. The influence of protein environment on the low temperature electronic spectroscopy of Zn-substituted cytochrome c. *J. Phys. Chem. B* **104,** 6932–6941 (2000).

38. Wachter, R. M., Elsiger, M. A., Kallio, K., Hanson, G. T. & Remington, S. J., Structural basis of spectral shifts in the yellow-emission variants of green fluorescent protein. *Struct. Folding and Design* **6,** 1267–1277 (1998).

39. Kummer, A. D. *et al.* Effects of threonine 203 replacements on excited-state dynamics and fluorescence properties of the green fluorescent protein (GFP). *J. Phys. Chem B* **104,** 4791–4798 (2000).

40. Lu, D., Chen, G., Perry, J. W. & Goddard, W. A. III. Valence-bond charge-transfer model for nonlinear optical properties of charge-transfer organic molecules. *J. Am. Chem. Soc.* **116,** 10679–10683 (1994).

41. Marder, S. R. *et al.* A unified description of linear and nonlinear polarization in organic polymethine dyes. *Science* **265,** 632–635 (1994).

42. Barzoukas, M., Runser, C., Fort, A. & Blanchard-Desce, M. A two-state description of (hyper)polarizabilities of push-pull molecules based on a two-form model. *Chem. Phys. Lett.* **257,** 531–537 (1996).

43. Drobizhev M. *et al.* Primary role of the chromophore bond length alternation in reversible photoconversion of red fluorescence proteins. *Sci. Rep.* **2,** 688 (2012).

44. Shin, Y.-G., Brunschwig, B. S., Creutz, C. & Sutin, N. Toward a quantitative understanding of dipole-moment changes in charge-transfer transitions: Electroabsorption spectroscopy of transition-metal complexes. *J. Am. Chem. Soc.* **117,** 8668–8669 (1995).

45. Asselberghs, I. *et al.* Second-harmonic generation in GFP-like proteins. *J. Am. Chem. Soc.* **130,** 15713–15719 (2008).

46. De Meulenaere, E. *et al.* Second-order nonlinear optical properties of fluorescent proteins for second-harmonic imaging. *J. Mater. Chem.* **19,** 7514–7519 (2009).

47. Brunschwig, B. S., Ehrenson, S. & Sutin, N. Solvent reorganization in optical and thermal electron-transfer processes: Solvatochromism and intramolecular electron-transfer barriers in spheroidal molecules. *J. Phys. Chem.* **91,** 4714–4723 (1987).

48. Fromherz, P. Monopole-dipole model for symmetrical solvatochromism of hemicianine dyes. *J. Phys. Chem.* **99,** 7188–7192 (1995).

49. Das, A., Hasegawa, J.-Y., Miyahara, T., Ehara, M. & Nakatsuji, H. Electronic excitations of the green fluorescent protein chromophore in its protonation states: SAC/SAC-CI study. *J. Comput. Chem.* **24,** 1421–1431 (2003).

50. Webber, N. M. & Meech, S. R. Electronic spectroscopy and solvatochromism in the chromophore of GFP and Y66F mutant. *Photochem. Photobiol. Sci.* **6,** 976–981 (2007).

51. Rebane, A. *et al.* Quantitative prediction of two-photon absorption cross section based on linear spectroscopic properties. *J. Phys. Chem. C* **112,** 7997–8004 (2008).

52. Bublitz, G., King, B. A. & Boxer, S. G. Electronic structure of the chromophore in green fluorescent protein (GFP). *J. Am. Chem. Soc.* **120,** 9370–9371 (1998).

53. Drobizhev, M., Makarov, N. S., Tillo, S. E., Hughes, T. E. & Rebane, A. Two-photon absorption properties of fluorescent proteins. *Nat. Meth.* **8,** 393–399 (2011).

54. Barnett, L. M. *et al.* The two-photon bazooka: A new way of optically screening randomly mutagenized libraries of fluorescent proteins. *Biophys. J.* **106,** Suppl. 1, 810A–811A (2014).

55. Dubinina, G. G. *et al.* Phenylene vinylene Platinum(II) acetylides with prodigious two-photon absorption. *J. Am. Chem. Soc.* **134,** 19346–19349 (2012).

56. Clark, T. B., Orr, M. E., Flynn, D. C. & Goodson T. III. Synthesis and optical properties of two-photon absorbing GFP-type probes. *J. Phys. Chem. C* **115,** 7331–7338 (2011).

57. Brooks, B. R. *et al.* CHARMM: The biomolecular simulation program. *J. Comput. Chem.* **30,** 1545–1614 (2009).

58. Ridley, J. & Zerner, M. Intermediate neglect of differential overlap (INDO) technique for spectroscopy: pyrrole and the azines. *Theor. Chim. Acta (Berl.)* **32,** 111–134 (1973).

59. Cory, M. G., Zerner, M. C., Xu, X. C. & Schulten, K. Electronic excitations in aggregates of bacteriochlorophylls. *J. Phys. Chem. B* **102,** 7640–7650 (1998).

60. Li, J., Williams, B., Cramer, C. J. & Truhlar, D. G. A class IV charge model for molecular excited states. *J. Phys. Chem.* **110,** 724–733 (1999).

61. Griesbeck, O., Baird, G. S., Campbell, R. E., Zacharias, D. A. & Tsien, R. Y. Reducing the environmental sensitivity of yellow fluorescent protein - Mechanism and applications. *J. Biol. Chem.,* **276,** 29188–29194 (2001).

62. Barstow, B., Ando, N., Kim, C. U. & Gruner, S. M. Alteration of citrine structure by hydrostatic pressure explains the accompanying spectral shift. *Proc. Natl. Acad. Sci. USA* **105,** 13362–13366 (2008).

63. Royant, A. & Noirclerc-Savoye, M. Stabilizing role of glutamic acid 222 in the structure of enhanced green fluorescent protein. *J. Struct. Biol.* **174,** 385–390 (2011).

64. Arpino, J. A. J., Rizkallah, P. J. & Jones, D. D. Crystal structure of enhanced green fluorescent protein to 1.35 Angstrom resolution reveals alternative conformations for Glu222. *PloS one* **7,** e47132 (2012).

65. Henderson, J. N., Ai, H. W., Campbell, R. E. & Remington, S. J. Structural basis for reversible photobleaching of a green fluorescent protein homologue. *Proc. Natl. Acad. Sci. USA* **104** (16), 6672–6677 (2007).

66. Ai, H. W., Henderson, J., Remington, S. & Campbell, R. Directed evolution of a monomeric, bright and photostable version of Clavularia cyan fluorescent protein: structural characterization and applications in fluorescence imaging. *Biochem. J.* **400,** 531–540 (2006).

67. Best, R. B. & Hummer, G. H. Optimized molecular dynamics force fields applied to the helix-coil transition of polypeptides. *J. Phys. Chem. B* **113,** 9004–9015 (2009).

68. Lindorff-Larsen, K. *et al.* Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins-Struct. Funct. Bioinformatics* **78,** 1950–1958 (2010).

69. Nifosì, R. & Tozzini, V. Molecular dynamics simulations of enhanced green fluorescent proteins: Effects of F64L, S65T and T203Y mutations on the ground-state proton equilibria. *Proteins-Struct. Funct. Genetics* **51,** 378–389 (2003).
70. Hess, B., Bekker, H., Berendsen, H. J. C. & Fraaije, J. G. E. M. LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **18,** 1463–1472 (1997).
71. Bussi, G., Donadio, D. & Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **126(1),** 014101 (2007).
72. Hess, B., Kutzner, C., van der Spoel, D. & Lindahl, E. GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* **4,** 435–447 (2008).
73. Aksimentiev, A. & Schulten, K. Imaging alpha-hemolysin with molecular dynamics: Ionic conductance, osmotic permeability, and the electrostatic potential map. *Biophys. J.,* **88,** 3745–3761 (2005).

Acknowledgments

Author Contributions

M.D. analyzed the experimental results, developed the physical model, and wrote the main manuscript with inputs from P.R.C., R.N., A.R., P.R.C. and R.N. performed quantum mechanical and M.D. simulations, G.W. and C.S. acquired experimental data, L.B. and T.E.H. designed, expressed and purified new G.F.P. mutants, P.S. synthesized and characterized the H.B.D.I. compound, A.R. designed the experiment. All authors reviewed the manuscript.

Additional Information

Supplementary information accompanies this paper at http://www.nature.com/srep

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Drobizhev, M. *et al.* Long- and Short-Range Electrostatic Fields in GFP Mutants: Implications for Spectral Tuning. *Sci. Rep.* **5**, 13223; doi: 10.1038/srep13223 (2015).