Genetically encoded fluorescent labels from the green fluorescent protein (GFP) family enable in vivo imaging of protein localization and interactions. In the context of biomedical applications, deep tissue imaging calls for proteins with absorption/emission above 600 nm (or below 2.07 eV). Most of the existing red fluorescent proteins (RFPs) share a DsRed-like chromophore in which the π-system of the GFP chromophore is extended by an additional N-acylimine moiety. One of the challenges in RFP engineering is their relatively low optical output and photobleaching. Lower quantum yield is attributed to an extended range of motion of the chromophore; thus, using a structural motif based on a smaller (and more rigid) chromophore may help to improve optical output of red FPs by reducing losses due to radiationless relaxation and photobleaching.

SECTION: Molecular Structure, Quantum Chemistry, and General Theory

A bathochromic shift can also be achieved by π-stacking interactions of the chromophore with two nearby aromatic rings (Phe, His, Trp, Tyr). One of the recent examples of in silico protein design has predicted red-shifted aromatic ring stacking in FPs from the GFP family. By introducing four point mutations, we arrive to a different structural motif based on a smaller chromophore that does not have this problem to begin with might be more fruitful. The authors applied the Thr to Tyr mutation, identical to that used in YFP, along with the modification of the residues around the acylimine oxygen and the phenolate moiety of the chromophore. The π-stacking motif has also been exploited in TagRFP56718 in which the red chromophore is stacked with histidine (mutation at position 69), yielding the excitation/emission maxima at 611/657 nm.

Here, we outline a strategy for increasing the bathochromic shift in the GFP-like proteins by further extending π-stacking interactions of the chromophore with two nearby aromatic amino acid residues. We computationally constructed GFP variants in which the conventional anionic GFP chromophore, 4-hydroxybenzylidene-imidazoline, is sandwiched between the two tyrosine residues, resulting in a "triple-decker" motif. To this end, we performed a few point mutations, in addition to the known modification Thr203Tyr4,15,16 The critical issue is to replace Thr62 by Tyr and to make additional changes in the protein design.

Among fluorescent proteins (FPs) used as genetically encoded fluorescent tags, the red-emitting FPs are of particular importance as suitable markers for deep tissue imaging. Using electronic structure calculations, we predict a new structural motif for achieving red-shifted absorption and emission in FPs from the GFP family. By introducing four point mutations, we arrive to the structure with the conventional anionic GFP chromophore sandwiched between two tyrosine residues. Contrary to the existing red FPs in which the red shift is due to extended conjugation of the chromophore, in the triple-decker motif, the chromophore is unmodified and the red shift is due to π-stacking interactions. The absorption/emission energies of the triple-decker FP are 2.25/2.16 eV, respectively, which amounts to shifts of ~40 (absorption) and ~25 nm (emission) relative to the parent species, the I form of wtGFP. Using a different structural motif based on a smaller chromophore may help to improve optical output of red FPs by reducing losses due to radiationless relaxation and photobleaching.

**ABSTRACT:** Among fluorescent proteins (FPs) used as genetically encoded fluorescent tags, the red-emitting FPs are of particular importance as suitable markers for deep tissue imaging. Using electronic structure calculations, we predict a new structural motif for achieving red-shifted absorption and emission in FPs from the GFP family. By introducing four point mutations, we arrive to the structure with the conventional anionic GFP chromophore sandwiched between two tyrosine residues. Contrary to the existing red FPs in which the red shift is due to extended conjugation of the chromophore, in the triple-decker motif, the chromophore is unmodified and the red shift is due to π-stacking interactions. The absorption/emission energies of the triple-decker FP are 2.25/2.16 eV, respectively, which amounts to shifts of ~40 (absorption) and ~25 nm (emission) relative to the parent species, the I form of wtGFP. Using a different structural motif based on a smaller chromophore may help to improve optical output of red FPs by reducing losses due to radiationless relaxation and photobleaching.

**SECTION:** Molecular Structure, Quantum Chemistry, and General Theory

A bathochromic shift can also be achieved by π-stacking interactions. For example, a replacement of the Thr203 residue in GFP by Tyr has led to the yellow FP (YFP) featuring a 20 nm red shift in absorption/emission relative to the parent species. As discussed in ref 15, the Thr203Tyr variant led to the most noticeable emission red shift among the mutants in which Thr203 was replaced by amino acid residues with planar aromatic rings (Phe, His, Trp, Tyr). One of the recent examples of in silico protein design has predicted red-shifted (up to 26 nm relative to the parent protein) mutants of mCherry; the theoretical predictions were confirmed experimentally. The authors applied the Thr to Tyr mutation, identical to that used in YFP, along with the modification of the residues around the acylimine oxygen and the phenolate moiety of the chromophore. The π-stacking motif has also been exploited in TagRFP56718 in which the red chromophore is stacked with histidine (mutation at position 69), yielding the excitation/emission maxima at 611/657 nm.

Here, we outline a strategy for increasing the bathochromic shift in the GFP-like proteins by further extending π-stacking interactions of the chromophore with two nearby aromatic amino acid residues. We computationally constructed GFP variants in which the conventional anionic GFP chromophore, 4-hydroxybenzylidene-imidazoline, is sandwiched between the two tyrosine residues, resulting in a "triple-decker" motif. To this end, we performed a few point mutations, in addition to the known modification Thr203Tyr. The critical issue is to replace Thr62 by Tyr and to make additional changes in the...
structure resulting in up to four modifications, which were required to achieve structures compliant with the Ramachandran plot.

The computational strategy used in this work was validated by computing structures and spectra of wtGFP with the anionic chromophore. Then the same approach was applied for modeling the mutated GFP with the triple-decker motif of the chromophore pocket. For full computational details, see the Supporting Information for this Letter. In brief, we used the flexible effective fragment version 19,20 of the quantum mechanics−molecular mechanics (QM/MM) approach 21 to optimize the equilibrium structures of the ground and excited states of the FP. The stationary points on the ground-state potential energy surface were located using QM(PBE0/6-31G*)/MM(AMBER). Optimization of the S1 excited-state geometry was performed by the QM(CASSCF(10/9)/6-31G*)/MM(AMBER) method by using complete active space self-consistent field (CASSCF). Calculations of the vertical $S_0 \rightarrow S_1$ and $S_1 \rightarrow S_0$ energy differences were performed using the extended multiconfigurational quasi-degenerate perturbation theory (XMCQDPT2) method, 22 which has been shown to yield accurate excitation energies in complex fluorophores (e.g., see ref 23). Further validation of the computed energies was carried out by another ab initio approach, the configuration interaction singles method augmented by perturbative doubles correction within the scaled-opposite-spin framework, SOS-CIS(D).24 As shown below, both approaches yield quantitatively consistent results.

A model system mimicking wtGFP was constructed by using the coordinates of heavy atoms from the crystal structure of the S65T mutant of GFP with the anionic chromophore (PDB id: 1EMA entry25). We restored the side chain of Ser at position 65 and added hydrogen atoms by assigning standard protonation states of polar residues, that is, positively charged Lys and Arg and negatively charged Glu and Asp.

Fragments of the QM/MM-optimized chromophore-containing domain corresponding to two known conformations of GFP with the anionic chromophore (the B and I forms26,27) are shown in Figure 1. The main difference between these forms is attributed to different conformations of the Glu222 side chain (the syn position in the B form and the anti position in the I form). The green fluorescence of wtGFP at 508 nm (2.44 eV) is assigned to the $S_1 \rightarrow S_0$ transition from the excited state of

Table 1. Wavelengths (nm, bold), Electronic Energy Differences (in parentheses, eV), and Computed Oscillator Strengths (in brackets) of the Electronic Transitions for the B and I Forms of wtGFP

<table>
<thead>
<tr>
<th>model structure</th>
<th>$S_0-S_1$ (absorption)</th>
<th>$S_1-S_0$ (emission)</th>
</tr>
</thead>
<tbody>
<tr>
<td>B form</td>
<td>XMCQDPT 484 (2.56)</td>
<td>SOS-CIS(D) 489 (2.53)</td>
</tr>
<tr>
<td></td>
<td>exp. 475 (2.61)</td>
<td>exp. 472 (2.62)</td>
</tr>
<tr>
<td></td>
<td>exp. 475 (2.50)</td>
<td>exp. 539 (2.30)</td>
</tr>
<tr>
<td>I form</td>
<td>XMCQDPT 493 (2.51)</td>
<td>SOS-CIS(D) 492 (2.52)</td>
</tr>
<tr>
<td></td>
<td>exp. 499 (2.62)</td>
<td>exp. 543 (2.28)</td>
</tr>
<tr>
<td></td>
<td>exp. 508 (2.44)</td>
<td>exp. 532 (2.30)</td>
</tr>
</tbody>
</table>

Figure 2. Fragment of the Thr62Tyr/Tyr145Ser/Ile167Gly/Thr203Tyr mutant showing all four replaced residues. Selected distances are given in Å; the bottom values for the Chr−Tyr203 pair refer to the crystal structure PDB id: 1YFP of the yellow mutant of GFP (Thr203Tyr/Ser65Gly/Val68Leu/Ser72Ala).
the I form, while the absorption at 475–472 nm (2.61–2.63 eV) is attributed to the B form, and the slightly redshifted band at 495 nm (2.51 eV) is attributed to the I form.28

The computed excitation energies and the corresponding experimental band maxima are collected in Table 1. First of all, we note an almost perfect correlation between the computed and experimental results for the absorption bands, which even reproduce subtle differences between the B and I forms. Slightly larger discrepancies are observed for emission bands; however, the differences (~0.1 eV) are within the established error bars of the QM and QM/MM protocols when applied for electronic transition energies of the protein-bound biochromophores.23,29–34

Below, we compare electronic transition energies of the mutated model protein with those of the I form of wtGFP (Figure 1, right panel).

Aiming to design a structure in which an almost planar chromophore is sandwiched between two planar π-systems, we manually prepared several structures by introducing four mutations in the wtGFP sequence. Our starting point was the manually prepared several structures by introducing four mutations Thr62Tyr/Thr203Tyr. To accommodate the Tyr side chain at position 62, we had to introduce a few more mutations; the minimal set of mutations in the wtGFP sequence. Our starting point was the mutated model protein with those of the I form of wtGFP (see Table 2).

Despite the slight variation in the computed values due to different approaches, the data in Table 2 clearly show that the mutant’s optical bands are red-shifted (relative to the I form of wtGFP) by ~0.2 (~40 nm) and ~0.1 eV (~25 nm) for absorption and emission, respectively, amounting to 2.25 eV absorption and 2.16 eV emission. At all levels of theory, we observe that mutations do not affect oscillator strengths significantly, and both transitions remain bright.

As discussed in the introductory part of this Letter, a stacked motif with a tyrosine near the chromophore has been exploited in several studies.15–17 A similar idea has been used in development of TagRFP56710,18 in which the red chromophore is stacked with histidine (mutation at position 69). One can also consider using another aromatic residue, Trp; however, its larger size could disrupt folding. We note that for practical applications properties beyond spectral shifts, such as brightness and photostability, are crucially important. While first-principle modeling of these properties is beyond the scope of this study, we note the experimental evidence16 that high quantum yields (e.g., >0.7 in some citrines) are attainable in π-stacked FPs.

In conclusion, a novel aspect of this study is a prediction of mutated structures in which the native GFP chromophore is sandwiched between two tyrosine residues. According to high-level calculations, these mutants featuring triple-decker stacked chromophores should exhibit noticeable red-shifted absorption and emission. The proposed motif can be used in directed mutagenesis studies aiming at engineering new RFP variants with improved optical properties.3–10 In particular, this motif can be used to achieve a large Stokes shift via the eximer-type relaxation, that is, without excited-state proton transfer that leads to lower quantum yields. Thus, our study presents a state structure was carried out using QM/MM. Figure 2 shows this model system with all four replaced residues. Note that the packing of the Chr–Thr203 pair in the computationally designed system resembles the corresponding motif in the crystal structure PDB id: 1YFP of the Thr203Tyr/Ser65Gly/Val68Leu/Ser72Ala GFP mutant.15,16

This model system is shown in Figure 3 from another angle allowing for direct comparison with wtGFP. We see that the most important features of the wtGFP machinery are not disturbed; the conservative Glu222 and Arg96 residues remain at their positions, and the hydrogen bond wire (Chr–Wat–Ser205–Glu222) is not disrupted.

The computed excitation energies for this mutant are collected in Table 2. The XMCQDPT2 values are computed with and without the field from the MM subsystem given in the columns marked “with EFP” and “no EFP”, respectively.

Table 2. Computed Wavelengths (nm, bold), Electronic Transitions Energies (eV, in parentheses), and Oscillator Strengths (in brackets) of the Triple-Decker Mutant of GFP and the Shifts of the Transition Energies Relative to the I Form of wtGFP (see Table 1)

<table>
<thead>
<tr>
<th></th>
<th>S0−S1 (with EFP)</th>
<th>S0−S1 (no EFP)</th>
<th>S0−S0 (with EFP)</th>
<th>S0−S0 (no EFP)</th>
</tr>
</thead>
<tbody>
<tr>
<td>wtGFP</td>
<td>[0.91] 527 (2.35) 550 (2.25)</td>
<td>530 (2.34)</td>
<td>[0.89] 555 (2.23) 572 (2.16)</td>
<td>569 (2.18)</td>
</tr>
<tr>
<td>shifts from wtGFP-I, eV</td>
<td>[0.97] 493 (2.51) 501 (2.47)</td>
<td>492 (2.51)</td>
<td>[0.87] 543 (2.28) 542 (2.28)</td>
<td>532 (2.33)</td>
</tr>
</tbody>
</table>

Figure 3. Model system with the triple-decker chromophore motif derived from wtGFP by introducing four mutations Thr62Tyr/Thr145Ser/Ile167Gly/Thr203Tyr.
starting point for a new route toward new RFPs featuring smaller chromophores in a tight environment.

ASSOCIATED CONTENT

Supporting Information
Detailed description of the computational protocols; Cartesian coordinates for the optimized structures; and relevant molecular orbitals. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author
*E-mail: anemukhin@yahoo.com; anem@lcc.chem.msu.ru.
Phone: +7-495-939-10-96.

Notes
The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This study was partially supported by the Program on Molecular and Cell Biology from the Russian Academy of Sciences and partly by the Russian Foundation for Basic Research (Project 13-03-00207). We acknowledge the use of supercomputer resources of the M.V. Lomonosov Moscow State University and of the Joint Supercomputer Center of the Russian Academy of Sciences. A.I.K. acknowledges the support from the National Science Foundation via the CHE-0951634 grant and from the Humboldt Foundation (Bessel Award).

REFERENCES

(31) Li, X.; Chung, L. W.; Mizuno, H.; Miyawaki, A.; Keiji Morokuma, K. A Theoretical Study on the Nature of On- and Off-States of Reversibly Photoswitching Fluorescent Protein Dronpa;

