Molecular Engineering of Torsional Potentials in Fluorogenic Dyes via Electronic Substituent Effects

Volkan Ediz, Jihoon L. Lee, Bruce A. Armitage, and David Yaron*

Department of Chemistry, Carnegie Mellon University, 4400 Fifth Avenue, Pittsburgh, Pennsylvania 15213

Received: June 23, 2008

Fluorogenic dyes such as thiazole orange (TO) and malachite green have been used in live cellular imaging due to their low quantum yield in solution but large fluorescence enhancements when bound to cellular nucleic acids or to a specific surface-expressed protein partner. Better understanding of the structure−property relationships that establish this fluorogenic behavior could benefit the design of improved dyes. In TO the fluorogenic properties are related to twisting of the dye, following electronic excitation in solution, from an emissive planar structure to a nonemissive twisted structure. Herein we develop a computational approach to identify electron acceptor/donor substitution patterns that impart desirable properties to the dye, such as inducing spectral shifts while maintaining an excited-state torsional surface that will lead to fluorogenic behavior. Additivity of substituent effects, on properties such as spectral shifts and excited-state torsional barriers, is tested and found to be sufficiently accurate that it can be used to identify promising dye candidates. Although additivity suggests an underlying linearity in the substituent effects, additional simplifications stemming from linearity could not be identified. The approach is tested on TO, considering seven different substituents at seven substitution positions, to identify fluorogenic dyes that will span a range of wavelengths. Additivity allows quantum chemical calculations on singly substituted molecules (49 molecules) to be used to make estimates for all substitution patterns (nearly 10^6 molecules).

1. Introduction

An important step toward understanding cellular signaling networks and pathways is monitoring cellular activity in real time.[1−3] Cellular imaging via fluoromodules holds great promise for providing such information.[1] A fluoromodule consists of a fluorescent dye and a specific binding element (e.g., an RNA, DNA or protein that binds specifically to the dye). Once an RNA or protein fluoromodule is developed, it can be attached to a biomolecule of interest by genetically fusing the sequence that codes for the binding element to the sequence that codes for the biomolecule. The binding elements that bind specifically to a chosen dye can be selected from combinatorial libraries, allowing us to focus here on optimization of the photophysical properties of the dye without regard to optimization of the binding.[4−6] Fluorogenic dyes are of particular interest because they have low quantum yield in solution but, when bound to the protein or RNA partner, large fluorescence enhancements are observed. This paper considers design of fluorogens for use in fluoromodules. The particular focus is on dye optimization through electron acceptor/donor substitution, because this is a particularly convenient means to manipulate the dye properties.[7]

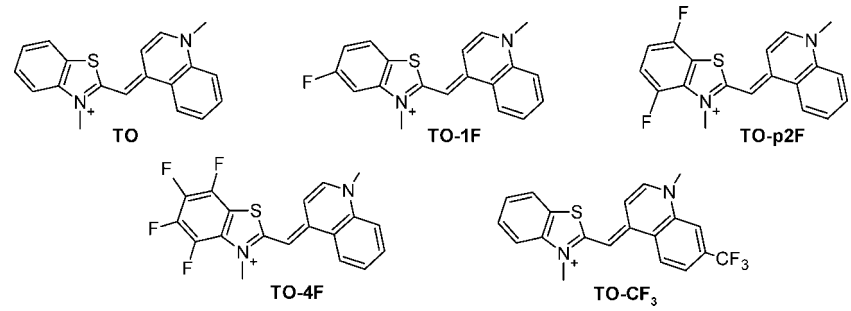
Design of dyes for fluoromodules is a challenging task because, in addition to satisfying the fluorogenic requirements, the dyes must satisfy a broad range of additional requirements such as specificity and photostability.[8] Cyanine dyes have been a particular focus in recent research and applications because, in addition to satisfying these requirements, they are sensitive fluorogens with high molar extinction coefficients and moderate quantum yields in solution.[9,10] Current fluorogenic cyanine dyes used as fluorescent labels include thiazole orange (**TO**) and oxazole yellow (**YO**).[11] Our goal is to use computation to identify candidate dyes that (i) span a wide range of wavelengths in the visible and near-IR range, so that multiple proteins may be monitored simultaneously in a living cell, and (ii) are highly fluorogenic. By remaining within the class of currently used cyanines, the dyes may continue to satisfy the additional requirements regarding photostability and specificity. This paper considers the use of electron acceptor/donor substituents to satisfy the two goals listed above. This double optimization problem has not yet been addressed by a systematic design approach. We are interested in filling this gap by developing an understanding of the structure−property relationships for fluorogenicity, along with strategies through which we can engineer improved fluorogenic properties (Chart 1).

Among the fluorogenic dyes, unsymmetrical cyanine dyes such as **TO** and **YO** are particularly attractive, because the excitation and emission wavelengths can be tuned throughout the visible range.[11] However, the effects of structural alterations on the fluorogenicity have not been studied in detail. Previously, fluorescence quenching in a small model cyanine and **TO** has been attributed to intramolecular twisting in the excited state.[12,13] Recently, we demonstrated that the fluorogenicity in **TO** originates from twisting of the dye from a planar structure that is highly fluorescent, to a nonplanar structure that is nonemissive.[14] As shown in Figure 1, there exists a small barrier between the planar local minimum associated with a large transition moment, and deep twisted minimum associated with near zero transition moment. The fluorogenic behavior results from the tendency of the dye, while in solution, to twist beyond a critical interplanar angle of 60° at which point it becomes dark. When bound in a conformationally constraining environment, such as the binding pocket of a protein or nucleic acid sequence this twisting is hindered and the dye becomes fluorescent. For a series of fluorinated **TO** analogues, the computed barrier heights

* Correspondence should be addressed to D.Y. (e-mail: yaron@cmu.edu).

10.1021/jp805546s CCC: $40.75 © 2008 American Chemical Society
Published on Web 09/11/2008

CHART 1: Structure of TO Dye and Its Fluorinated Derivatives



are well correlated with observed quantum yields in viscous solution.[14] This suggests that activated crossing of this barrier plays an important role in establishing the quantum yield in solution and hence the fluorogenic properties. In constrained environments, the substituent effects on quantum yield follow a different trend, suggesting that the quantum yield is established by a different mechanism such as coupling to stretching motions.[15] It also suggests that engineering the characteristics of the barrier is a useful design strategy for developing improved dyes. Consider, for instance, that changes made to the dye to shift the absorption wavelength, or optimize some other property, may raise the excited-state barrier to a point where the dye fluoresces even in solution, thereby destroying the fluorogenicity.

Here, we are interested in exploring the role of electronic substituent effects on the excited-state torsional potential by generating a library of **TO** analogues that vary in the type and position of electron acceptor/donor groups. For this purpose, we choose seven simple substituents, and consider eight positions on the heterocycles where these can be placed, as shown in Figure 2. In principle, this setup leads to a large database of nearly 10^6 molecules, which makes an exhaustive computational search infeasible. Thus, we need a systematic design approach that can reduce the computational cost of the search. Below, we show that additivity of substituent effects is a reasonable assumption for this system. Additivity allows calculations on singly substituted molecules (56 molecules in this case) to be used to make estimates for all substitution patterns (nearly 10^6 molecules). This allows us to simultaneously optimize multiple photophysical properties, as discussed above.

Additivity of the response from various substitutions is one of the consequences of linearity and so the additivity found here suggests that the system is in a regime where linear response theory holds. In addition to additivity, linearity implies that the response is proportional to the strength of the electronic substituents as captured by, for instance, a set of Hammett-like parameters. This latter aspect of linear response theory leads to consequences beyond additivity, none of which appear to be operable in the current situation. Consider, for instance, the internal field model of conjugated systems such as push−pull aromatic systems[16,17] and polyenes.[18,19] A consequence of this model is that two different properties, the hyperpolarizability and induced π electron dipole moment, are proportional to the strength of the acceptor−donor substitution pattern, as modeled by the magnitude of an internal electric field, **F**. Because both of these properties are proportional to **F**, they are correlated with one another and such correlations have been confirmed experimentally.[16,17] Weibel et al. generalized this approach to arbitrary substitution patterns by adopting a more general linear response theory.[20] When applied to **TO**, this generalized linear response theory would describe the effects of the substituents through a vector **V**. Each element of **V** corresponds to a particular position on the conjugated framework (P_1-P_8 shown in Figure 2) with a numeric value that describes the strength of the substituent at each position. The change in a property P induced by the substitution, ΔP, is then related to **V** through a general linear relation:

$$\Delta P = \mathbf{AV} \tag{1}$$

where **A** is a linear response matrix. As mentioned above, a linear relation leads to additivity of effects from different substituent patterns, V_1 and V_2:

$$\Delta P = A(V_1 + V_2) = AV_1 + AV_2 = \Delta P_1 + P_2 \tag{2}$$

We will show below that additivity does indeed apply for the properties of interest here. Equation 1 also implies that the substituent effects are proportional to some Hammett-like parameters, the elements of **V**, that capture the strength of the substituents. If this was the case, it would be possible to factor the responses ΔP into the product of a parameter that captures the strength of the substituent, the **V** of eq 1, and a position dependent proportionality constant, the **A** of eq 2. We will show below, however, that factorization of the form of eq 2 does not apply for the properties of interest here, and so this aspect of linearity does not apply.

The lack of linearity with respect to acceptor/donor strength is, perhaps, not surprising given past work on the effects of electronic substituents of the nonlinear optical properties of cyanines. This work found that the nonlinear optical susceptibilities could be understood in terms of a bond-length alternation (BLA) coordinate.[18,21] BLA is typically defined as the average difference between the single and double bond lengths along the relevant conjugation pathway. Electronic substituents may be expected to alter the relative importance of the two resonance structures of **TO** shown in Figure 2. These structures have opposite BLA and so the observed BLA may provide a useful measure of the relative importance of these two structures. The nonlinear optical susceptibilities follow a regular behavior versus BLA; however, this behavior is far from linear. BLA may therefore provide a means to go beyond the linear form of eq 2 and better understand the trends resulting from different substituents. Below, we will show that many properties of **TO** do indeed show good correlations with BLA; however, the excited-state barrier, the key property for fluorogenicity, is not well correlated with the BLA. Unlike the case of nonlinear optical susceptibilities, the BLA does not provide a useful way to understand the relation between the excited-state barrier height and the acceptor−donor strength. It is fortunate, however, that such an understanding is not necessary to support computational design of improved fluorogens. Additivity alone is sufficient to allow a manageable number of quantum chemical calculations (56 calculations corresponding to singly substituted molecules)

to be used to enable an exhaustive search of all substitution patterns (nearly 10^6 molecules) on the **TO** conjugated framework.

The choice of chemical substituents and computational methods are described in section 2, including a comparison of the semiempirical quantum chemical methods used here with results from time-dependent density functional theory (TD-DFT).[22] We begin, in section 3.1, by using BLA as a descriptor for the structure−property relationships. First, we present a principal component analysis (PCA) of all possible BLA coordinates, as shown in Figure 2, and then investigate how key components of the excited-state torsional potential correlate with the primary BLA coordinate. In section 3.2, we show that an additivity model derived from singly substituted molecules can be used as a rational design tool for improved cyanine dyes. In section 3.3, we show that, despite the additivity found in section 3.2, linearity is not useful for understanding the relation between properties and substituent strength. Finally, in section 3.4, we will evaluate our design approach on two test cases.

2. Methods

2.1. Model. Molecular structures were generated using an automated system that creates initial geometries for a series of **TO** analogues by replacing hydrogen atoms on the two heterocycles with the desired substituents. This system is capable of generating multiply substituted analogues, allowing us to create a large library of compounds which we can then mine to derive structure−property relationships. To explore the effects of substitutions on the photophysical properties, seven substituents ($N(CH_3)_2$, NH_2, CH_3, F, CHF_2, CF_3, CN), were chosen that span a range of Hammett parameters, from −0.8 to +0.8[23] in a fairly evenly distributed manner. This ensures that the database is not biased toward a particular electron acceptor or donor strength. The eight substitution positions considered are shown in Figure 2 and denoted as P_i, where i is the index of the position. These substitution positions do not interfere with the main conjugation path between the nitrogen atoms on the heterocycles, such that the key components of the excited-state torsional potential, and hence fluorogenic properties, are likely to be preserved. A database of 1024 molecular structures was created including all single and double substitutions in which at most one group is present on each heterocycle. However, as we will discuss further below, substitution at P_5 introduces steric interactions that substantially alter the shape of the torsional potential. Therefore, molecular structures that involve substitution at P_5 were removed from the data. This reduced our database to 768 molecular structures, corresponding to placing seven electron donor/acceptor groups at seven positions.

Molecular structures were optimized using the SAM1[24] Hamiltonian in the AMPAC[25] quantum chemistry package. The ground electronic state was optimized using restricted Hartree−Fock (RHF) theory. Although we use SAM1 for the geometry optimization, we previously demonstrated that the vertical excitation energies predicted by SAM1 do not correlate well with experiments for fluorinated **TO** analogues.[14] More reliable predictions of the excitation energy can be obtained from Intermediate Neglect of Differential Overlap (INDO) theory.[26] However, because INDO is only useful after the molecular geometry is known, we adopt a hybrid method in which geometries are obtained from SAM1 and excitation energies are obtained from a direct INDO/SCI[27] method that includes excitations between all molecular orbitals. As shown in section 2.2, this hybrid method leads to good correlation between theory and experiment for both absorption and fluorescence (data not shown here), supporting the use of this approach to study substituent effects on the excited-state torsional potential.

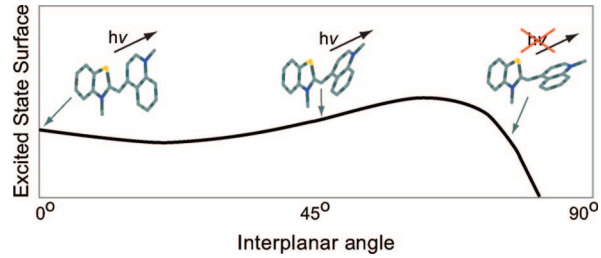


Figure 1. Excited-state energy surface of **TO** as a function of interplanar angle.

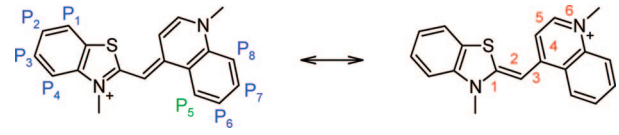


Figure 2. Resonance structures of **TO**. Substitution positions, P_i, are shown in blue and green on the left resonance structure. The bonds along the main conjugation pathway are enumerated in red on the right structure.

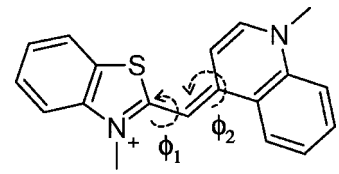


Figure 3. Definition of dihedral angles, ϕ_1 and ϕ_2, used to describe twisting of **TO**.

Figure 3 defines two interplanar angles ϕ_1 and ϕ_2 that can be used to characterize the twisting motion in **TO**. However, only twisting about ϕ_1 was examined because we previously showed that the reaction path is primarily along ϕ_1.[14] To reconfirm this, we randomly chose 20 molecular structures substituted with strong electron donors/acceptors at various positions and investigated whether the reaction path changes to involve twisting about ϕ_2. For all structures, the reaction path was along ϕ_1 (data not shown here), justifying the use of ϕ_1. The excited-state torsional surface is obtained by optimizing the ground-state geometry with constraints applied to hold ϕ_1 at values ranging from 0° to 90° in steps of 10°. The lowest singlet excitation energies were obtained from direct INDO/SCI calculations at each of these geometries.

2.2. Comparison of INDO/SCI with TDDFT. In an attempt to evaluate the performance of our hybrid method, we compared the predictions with those of time-dependent density functional theory (TDDFT), because TDDFT has had considerable success in predicting excitation energies.[28−30] We chose to perform calculations on the substituted **TO** molecules shown in Chart 1 because experimental data is available against which to test the methods.[14] Table 1 lists the excitation energies corresponding to the absorption maxima calculated using four different hybrid approaches. The approaches pair two different methods of ground-state optimization (SAM1 and DFT) with two different methods for vertical excitation energies (TDDFT and INDO/SCI). The methods used for SAM1 geometry optimization and INDO/SCI excitation energies are described in section 2.1. The DFT ground-state geometries were fully optimized using the hybrid B3YLP exchange-correlation functional and the 6-311G* basis set. A tight convergence threshold of 1.5×10^{-5} hartree/bohr was employed. The TDDFT vertical excitation energies were computed using the B3YLP functional and 6-311G* basis set. All calculations at the DFT level were carried out with the GAUSSIAN03 software package.[31]

TABLE 1: Results from Modeling Electronic Spectra of TO and Fluorinated Analogues

dye	experiment λ_{max} (nm)[a]	B3YLP/TDDFT λ_{max} (nm)[b]	B3YLP/INDO λ_{max} (nm)[c]	SAM1/INDO λ_{max} (nm)[d]	SAM1/TDDFT λ_{max} (nm)[e]
TO	502	432.4	465.6	476.5	445.1
TO-1F	500	436.6	461.5	471.4	449.7
TO-p2F	489	426.2	457.9	468.3	440.2
TO-4F	485	430.9	452.7	463.0	450.5
TO-CF3	516	437.7	470.1	482.5	450.5

[a] Experimental absorption maximum in methanol. [b] Excitation energy from B3YLP/6-311G*/TDDFT for B3YLP/6-311G* ground-state optimized structure. [c] Excitation energy from direct INDO/SCI for B3YLP/6-311G* ground-state optimized structure. [d] Excitation energy from direct INDO/SCI for SAM1 ground-state optimized structure. [e] Excitation energy from B3YLP/6-311G*/TDDFT for SAM1 ground-state optimized structure.

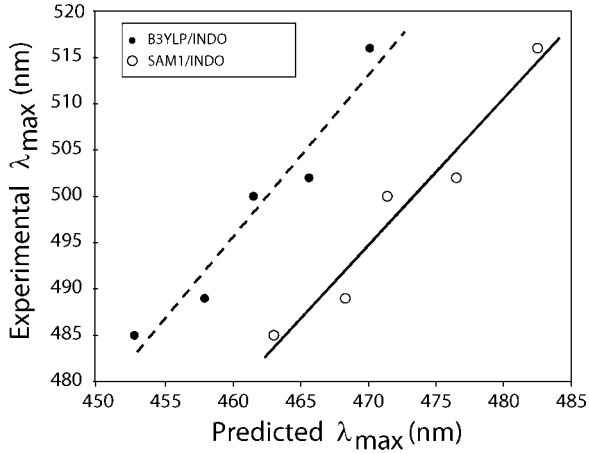


Figure 4. Correlation of predicted λ_{max} with experimental λ_{max}, in nm for the molecules of Chart 1. The dashed line is for B3YLP/6-311G*/INDO, and the solid line is for SAM1/INDO.

The results in Table 1 and Figure 4 show that the vertical excitation energies predicted by TDDFT do not correlate well with the experiment, regardless of choice of the geometry optimization method. This is surprising, because Champagne et al. found that the TDDFT scheme performed well for the optical properties of a set of symmetrical cyanine dyes.[32] The excitation energies predicted from INDO/SCI strongly correlates with the experiment regardless of the optimization method, indicating the INDO/SCI calculations are useful to understand the effects of substitution on **TO**. We note that the absolute magnitudes are not as reliable as the correlation, because a fit of experiment to theory for absorption yields a slope of 1.58 for INDO on SAM1 geometries and 1.75 for INDO on B3YLP geometries. The value of r^2 is 0.94 in both cases, indicating equal predictive power is obtained using SAM1 or B3YLP to optimize the geometry. Because the SAM1 method substantially reduces the computational cost, we chose to perform all our calculations using this hybrid approach.

3. Results

3.1. PCA and Use of BLA as a Coordinate. Figure 2 labels six carbon−carbon bonds along the main conjugation pathway between the nitrogen atoms on the heterocycles. Feature extraction algorithms provide a way to determine if variation in the length of these bonds between various substitution patterns can be well described by a single, or few, bond length alternation coordinates.[33] In this case, the features are the individual bond lengths and the resulting principle components capture the degree to which these bond lengths vary in a concerted manner across various substitution patterns.

Principal component analysis (PCA) is a widely used feature extraction method because of its simplicity and generality.[33]

TABLE 2: Results of PCA on the Bond Lengths for Singly and Doubly Substituted TO Analogues[a]

	bond lengths						
	r_1	r_2	r_3	r_4	r_5	r_6	% variance
PC1	−0.41	0.53	−0.55	0.33	−0.27	0.23	77.5
PC2	−0.24	0.27	−0.31	−0.52	0.47	−0.51	14.4
PC3	0.03	0.16	0.06	−0.77	−0.43	0.41	4.5

[a] Coefficients of the first three principal components are given along with the percent variance explained by each principal component.

PCA involves an orthogonal linear transformation of the feature space in such a way that the first linear combination of the original features, the first principal component, explains the greatest variance in data, the second principal component is orthogonal to the first one and explains the greatest remaining variance and so on. Thus, each principal component identifies and ranks the most important features needed to capture the variability in data. Principal components extracted in this way define a new feature space that contains the same information as the original feature space but with reduced dimensionality. Typically, only a few principal components are sufficient to capture the variability in data.

The results of a PCA analysis for the six bond lengths of Figure 2, performed on the database of singly and doubly substituted molecules (see section 2.1) are shown in Table 2. The results show that the first two principal components account for over 90% of the total variability. The first component corresponds to the change in bond lengths expected for a transition between the two resonance structures of Figure 2. The positive values of r_2, r_4, and r_6 indicate that the change in the bond lengths is in the same direction, and negative values of r_1, r_3, and r_5 indicate that the change in the bond lengths is in the opposite direction. The magnitude of the coefficients of r_2 and r_3 indicate that the largest variability in bond lengths is on the monomethine bridge. This, along with the large percent variance accounted for by a single principal component, justifies the use of BLA on the methine bridge ($r_3 - r_2$) as a measure of substituent effects. The second and the third principal component represent asymmetric bond distortions, with large coefficients associated with r_4, r_5, and r_6. This indicates that the bond lengths on the quinoline ring are more sensitive to substitution than those on the benzothiazole ring. Nevertheless, the percent variance explained by these principle components is small compared to that of first principal component, supporting the use of a single BLA coordinate such as ($r_3 - r_2$).

We next consider whether the substituent effects on the excited-state torsional potential can be understood in terms of the BLA coordinate on the methine bridge ($r_3 - r_2$) of Figure 2. Because our hybrid method obtains the excited-state torsional potential as the sum of the ground-state energy and the lowest

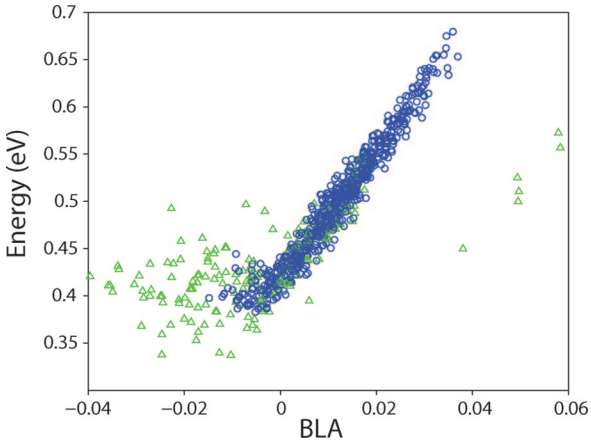


Figure 5. Correlation of the barrier to rotation in the ground state with the bond length alternation. The barrier is defined as the difference in energy between the planar and twisted ($\phi_1 = 90°$) structures. Molecular structures having substituents at P_5 are shown as green triangles, and structures with any other substitution pattern are shown as blue circles (see Figure 2).

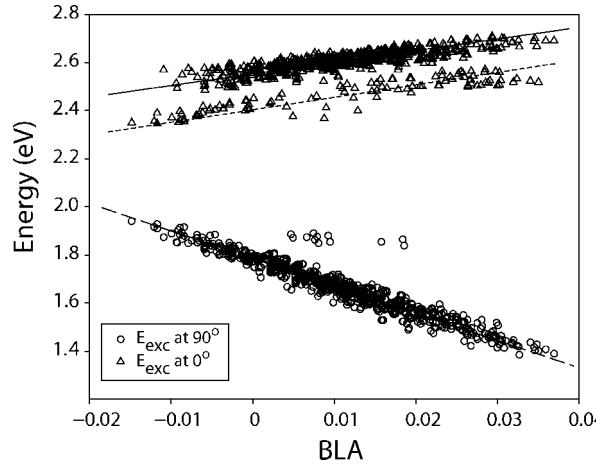


Figure 6. Correlation of the lowest singlet excitation energy with the bond length alternation. The excitation energies of the planar structure (ϕ_1 fixed at 0°) are shown as triangles, and energies for the twisted structure (ϕ_1 fixed at 90°) are shown as circles. The upper, solid, line shows a linear fit for the excitation energy of the planar structures excluding a subgroup in which strong electron donors ($N(CH_3)_2$ or NH_2) are substituted at P_2. The middle, dotted, line shows a linear fit for the excitation energy of the excluded subgroup. The lower, dashed, line shows a linear fit of the excitation energy of the twisted structures for all substitution patterns.

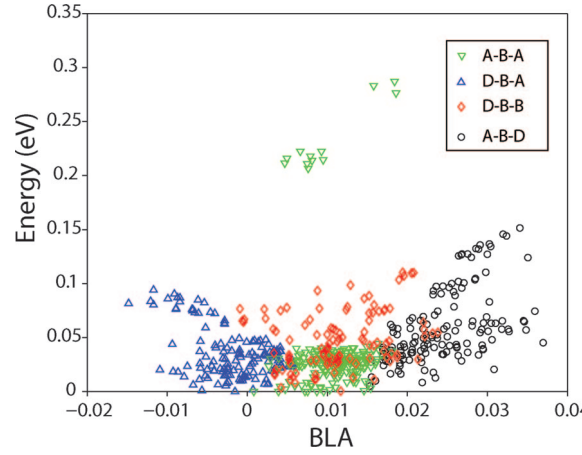


Figure 7. Correlation of the excited-state barrier with bond length alternation. The barriers of molecules with acceptor−bridge−donor (A−B−D) motif are shown as black circles, and barriers of donor−bridge−acceptor (D−B−A), acceptor−bridge−acceptor (A−B−A), and donor−bridge−donor (D−B−D) motifs are shown as blue upper triangles, green lower triangles, and red diamonds, respectively.

singlet excitation energy, we first consider separately the correlation of the ground-state energy and excitation energy with the BLA coordinate. The barrier to rotation in the ground electronic state is shown in Figure 5. The barrier correlates well with the BLA coordinate for substitution at all positions except P_5. Substituents at P_5 sterically interact with the hydrogen atom on the bridge when the structure is planar. As the structure twists along ϕ_1, these steric interactions weaken so that the ground-state rotational barrier drops substantially. Because the torsional potential is strongly influenced by steric interactions, and our goal here is to understand electronic substituent effects, the molecular structures that include substitution at P_5 are eliminated from the library of **TO** derivatives for the remainder of this analysis. For all other structures, a strong correlation of barrier and BLA is observed, implying that BLA is a useful parameter for understanding this feature of the torsional potential.

Our hybrid method obtains the excited-state torsional potential as a sum of the ground-state potential and excitation energy. Above, we found that the ground-state rotational barrier correlates well with BLA. We next consider the excitation energy. For all molecules, the excitation energy drops monotonically between $\phi_1 = 0°$ and $\phi_1 = 90°$. Figure 6 shows the INDO excitation energies at $\phi_1 = 0°$ and $\phi_1 = 90°$ as a function of the BLA coordinate for the substituted **TO** dyes. The results show that the BLA coordinate and the excitation energies at $\phi_1 = 0°$ and $\phi_1 = 90°$ are well correlated, although with opposite slopes for the correlation. A subgroup of **TO** derivatives, in which $N(CH_3)_2$ or NH_2 are substituted at P_2, leads to a systematically lower excitation energy at $\phi_1 = 0°$. The best fit line of the excitation energy for this subgroup has a slope similar to that of the other structures. This behavior suggests a stronger perturbation of the π electrons on the benzothiazole ring for such substitutions, as is further confirmed in section 3.2. Another subgroup, in which CF_3 is substituted on benzothiazole ring, and CHF_2 is substituted on the quinoline ring, leads to set of outliers seen for the excitation energy at $\phi_1 = 90°$. The reasons that these substitutions lead to outliers are not clear, however, we will see below that the outlier character of these structures applies also to other properties. The number of molecules in these two subgroups is quite small compared to that of the full database, thus the BLA coordinate can still be taken as having a reasonable correlation with the excitation energies at $\phi_1 = 0°$ and $\phi_1 = 90°$.

The feature of the excited-state torsional potential that likely establishes the fluorogenic mechanism in these dyes is the barrier to rotation from the planar, emissive, structure to the twisted, nonemissive, structure. The excited-state surface was calculated as the sum of the ground-state surface and excitation energy, both of which were shown above to be reasonably well correlated with the BLA coordinate. However, as shown in Figure 7, a similarly simple correlation does not exist for the barrier between the planar and the twisted configurations. The results are shown for four separate substitution motifs in which at most one acceptor/donor group is present on each heterocycle: donor on the benzothiazole−bridge−donor on the quinoline ring (D−B−D), acceptor−bridge−donor (A−B−D), donor−bridge−acceptor (D−B−A), and acceptor−bridge−acceptor (A−B−A) compounds. Decomposing the excitation energy into contributions from orbital ($E_{HOMO} - E_{LUMO}$), Coulomb (J) and exchange

TABLE 3: Changes in the Excited-State Barrier Height Induced by Single Substitutions on the TO Framework (eV)[a]

	P_1	P_2	P_3	P_4	P_6	P_7	P_8
$N(CH_3)_2$	−0.0111	0.0392	0.0029	−0.0292	0.0843	0.0287	0.0200
NH_2	−0.0067	0.0404	0.0060	−0.0167	0.0517	0.0063	0.0262
CH_3	−0.0029	0.0108	−0.0012	−0.0196	0.0080	0.0004	0.0038
F	0.0033	0.0038	0.0009	−0.0010	0.0087	−0.0018	0.0034
CHF_2	0.0002	0.0025	−0.0044	−0.0243	0.0085	−0.0002	−0.0016
CF_3	0.0056	0.0081	−0.0007	−0.0194	−0.0085	−0.0009	−0.0064
CN	−0.0003	0.0050	−0.0057	−0.0132	0.0093	0.0064	0.0043

[a] The substitution positions are denoted as shown in Figure 2.

TABLE 4: Changes in the Vertical Excitation Energy (ϕ_1 Fixed at 0°) Induced by Single Substitutions on the TO Framework (eV)[a]

	P_1	P_2	P_3	P_4	P_6	P_7	P_8
$N(CH_3)_2$	−0.0140	−0.2119	−0.0562	−0.0349	−0.1147	0.0262	−0.0026
NH_2	−0.0052	−0.1782	−0.0371	−0.0808	−0.0850	0.0341	−0.0319
CH_3	−0.0029	−0.0483	−0.0063	−0.0253	−0.0141	0.0152	−0.0059
F	0.0368	−0.0046	0.0276	0.0104	−0.0412	−0.0122	−0.0252
CHF_2	0.0345	0.0134	0.0205	0.0158	−0.0316	−0.0129	−0.0218
CF_3	0.0467	0.0409	0.0414	0.0392	−0.0404	−0.0330	−0.0320
CN	0.0393	0.0147	0.0350	0.0164	−0.0485	−0.0397	−0.0397

[a] The substitution positions are denoted as shown in Figure 2.

(K) terms did not shed further light on the origins of this poor correlation. All three of these contributions to the excited-state barriers are poorly correlated with the BLA (data shown in Supporting Information).

This prompted us to consider other factors that may play a role in establishing the barrier height. Analysis of the database reveals that this poor correlation may be attributed to the insensitivity of the BLA coordinate to the substitution position. Consider, for instance, molecules with CF_3 substitution at any position on the benzothiazole ring, and any other substituent on the quinoline ring (excluding P_5 as discussed above). For these molecules, the BLA is relatively independent of the specific position of the substituents whereas the excited-state barrier changes substantially with position (see Supporting Information). Similar behavior is seen for other substituent pairs. This indicates that factors that do not influence the BLA, such as specific location of substituents, do influence the excited-state barrier. We therefore need to go beyond BLA to understand the excited-state barrier, to a model that can distinguish between substitutions at various positions. This is analogous to the situation that arises when extending internal field models from unidirectional to multipolar molecules, for which past work found additivity to be a useful approach.[20]

3.2. Additivity Arguments. In this section, we consider whether the additivity of substituent effects on **TO** derivatives is a reasonable assumption. Consider, for instance, an additivity approximation for doubly substituted structures, which can be defined as

$$(P_{X-B-Y} - P_{H-B-H}) \approx (P_{X-B-H} - P_{H-B-H}) + (P_{H-B-Y} - P_{H-B-H}) \quad (3)$$

where P is the molecular property, B denotes the monomethine bridge, H denotes the hydrogen atom of an unsubstituted heterocycle, X denotes a substituent at a specific position of the benzothiazole ring, and Y denotes a substituent at a specific position on the quinoline ring. Tables 3 and 4 list the changes in excitation energy and the barrier height of the excited-state surface induced by single substitutions on the **TO** framework, as appear on the right-hand-side of eq 3. Response tables for other molecular properties are given as Supporting Information.

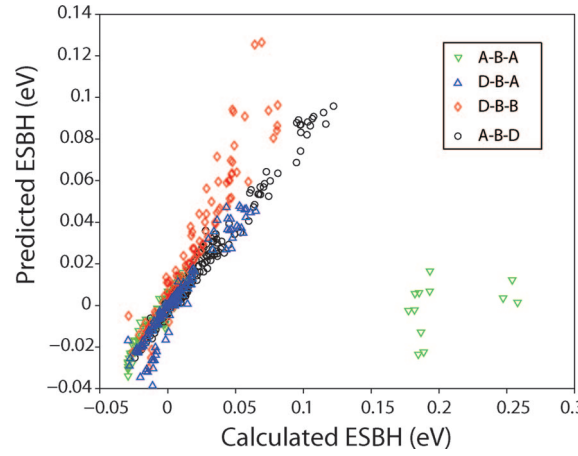


Figure 8. Correlation of the excited-state barrier height (ESBH), predicted assuming additivity, with the explicitly calculated ESBH for doubly substituted **TO** analogues.

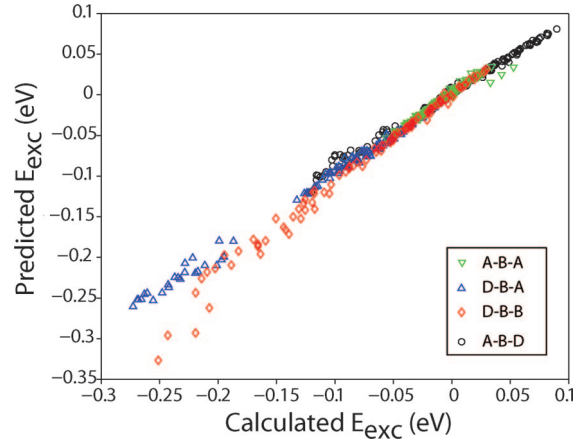


Figure 9. Correlation of the vertical excitation energy at the planar structure (ϕ_1 fixed at 0°), predicted assuming additivity, with the explicitly calculated vertical excitation energy for doubly substituted **TO** analogues.

Because we are interested in fluorogenic properties, we next consider whether additivity applies to the barrier height of the excited-state surface. Figure 8 shows that the additivity assumption holds for A−B−D, D−B−A, D−B−D compounds for which the r^2 values are 0.98, 0.91, and 0.90 and the slopes are 0.84, 0.87, and 1.23 respectively. For A−B−A compounds, as in Figure 6, we again find the CF_3−B−CHF_2 subgroup of molecules to be substantial outliers on an otherwise strongly correlation. When this outlying group is excluded, the value of r^2 is 0.92 with a slope of 0.94 for the A−B−A motif. To maintain fluorogenic behavior, the barrier must be in a range wherein the molecule will twist easily from a planar to a nonplanar structure in solution. The correlations found here suggest that the accuracy resulting from the additivity assumption is sufficient for the purposes of identifying dye candidates whose excited-state barriers are in a range where fluorogenic behavior can be expected.

Figure 9 shows that additivity also applies to the absorption wavelength for A−B−D, D−B−A, A−B−A and D−B−D compounds for which r^2 values are 0.99, 0.99, 0.97 and 0.98 and the slopes are 0.89, 0.95, 0.95, and 1.14, respectively. These correlations indicate that additivity provides a useful estimate of the spectral shift.

Figure 9 shows that a subclass of D–B–D molecules (red diamonds in lower left corner of the figure) shows the largest deviations from linearity. These are molecules in which $N(CH_3)_2$

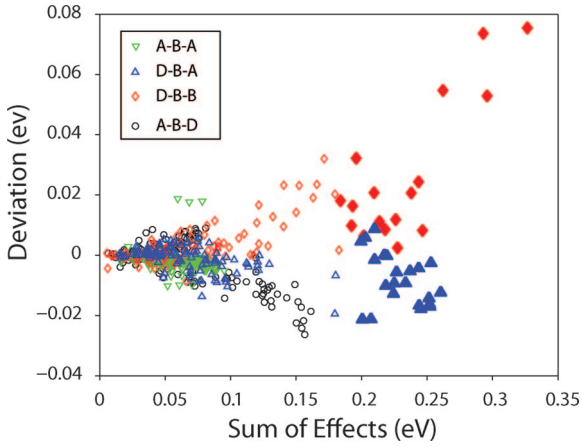


Figure 10. Plot of deviation from additivity against the sum of the absolute values of the shifts induced by the individual substituents. The results suggest the error associated with additivity increases with the combined strength of the substituents. The bolds symbols highlights the subgroup in which strong electron donors ($N(CH_3)_2$ and NH_2) are substituted at P_2.

or NH_2 are substituted at P_2. These stronger deviations from linearity may be related to these substitutions having the largest effects of any single substitutions. (Tables 3 and 4 show that substituting $N(CH_3)_2$ or NH_2 at P_2 leads to the largest effects on both the excitation energy and excited-state surface barrier height.) This suggests that deviations from linearity may be larger for stronger substituent effects. Figure 10 plots the deviation from linearity against the combined strength of the substitutions, as measured by the sum of the absolute values of the shifts induced by the individual substituents (Tables 3 and 4). The results support the notion that the deviation from linearity increases with combined substituent strength. From this perspective, the compounds with $N(CH_3)_2$ or NH_2 substitutions at P_2 deviate most strongly from linearity because they have the strongest combined substituent strength (bold symbols in Figure 10). The behavior in Figure 6, in which $N(CH_3)_2$ or NH_2 at P_2 lie on a line below that of the other molecules, can then be rationalized in terms of these strong substitutions leading to a red (bathochromic) shift that can not be accounted for by a generic correlation of excitation energy with BLA.

The above shows that additivity applies when substituents are placed on different heterocycles. We next test whether additivity also holds for double substitutions on the same heterocycle, for where the substituents are in much closer proximity. A representative set of 98 molecules was generated by considering all possible combinations of the substituent groups placed either at P_1 and P_3 on the benzothiazole ring or at P_6 and P_8 on the quinoline ring. These positions were chosen to avoid steric interactions both between the substituents and between the substituents and the bridge. Figures 11 and 12 indicate that additivity holds well for this set of molecules. Figure 11 plots the excitation energy obtained from explicit calculations versus that obtained assuming additivity. The value of r^2 is 0.98 and the slope is 0.98. Figure 12 is a similar plot for the excited-state barrier, for which r^2 is 0.94 and the slope is 0.85. Similarly good correlations are found for the other properties of interest (see Supporting Information). These results suggest that additivity applies even for substitutions that perturb the same heterocycle.

To test the additivity assumption further, we next considered triple and quadruple substitutions. A representative set of 40 molecules was generated by randomly selecting both the positions (excluding P_5 as discussed in section 3.1) and types

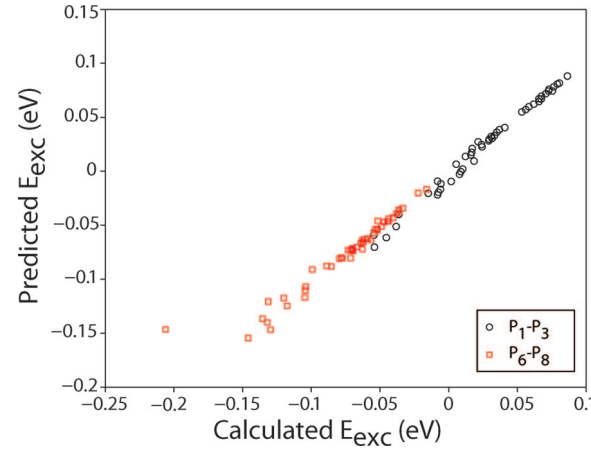


Figure 11. Correlation of the vertical excitation energy at the planar structure (ϕ_1 fixed at 0°), predicted assuming additivity, with the explicitly calculated vertical excitation energy for doubly substituted **TO** analogues involving substitution on the same heterocycle.

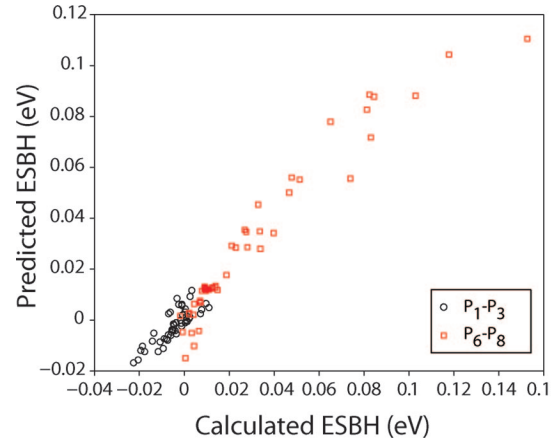


Figure 12. Correlation of the excited-state barrier height (ESBH), predicted assuming additivity, with the explicitly calculated ESBH for doubly substituted **TO** analogues involving substitution on the same heterocycle.

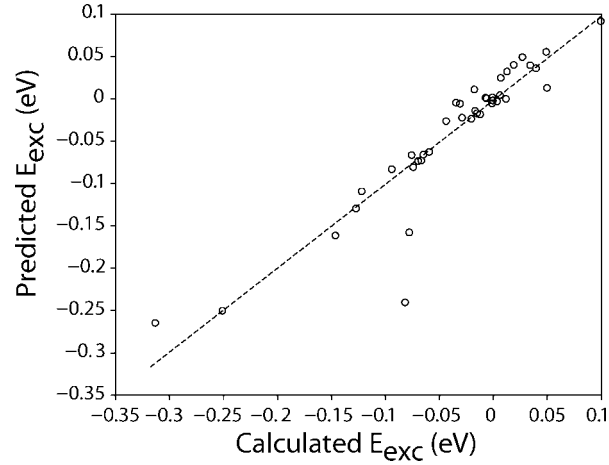


Figure 13. Correlation of the excitation energy (ϕ_1 fixed at 0°), predicted assuming additivity, with the explicitly calculated vertical excitation energy at the planar structure for triply and quadruply substituted **TO** analogues.

of the substituents. The molecules and results are listed in the Supporting Information. Figures 13 and 14 indicate that additivity holds for these molecules. Figure 13 plots the excitation energy obtained from explicit calculations versus that obtained assuming additivity. The value of r^2 is 0.84 and the slope is

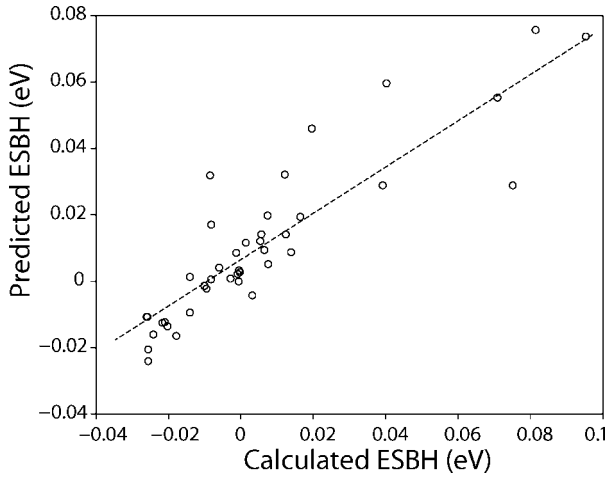


Figure 14. Correlation of the excited-state barrier height (ESBH), predicted assuming additivity, with the explicitly calculated ESBH for triply and quadruply substituted **TO** analogues.

0.99. Figure 14 is a similar plot for the excited-state barrier, for which r^2 is 0.75 and the slope is 0.70. Two of the randomly generated molecules have an $N(CH_3)_2$ or NH_2 at position P_2 and so, as discussed above, are expected to deviate most strongly from linearity. When these are removed from the data set, the correlations improve substantially, with the excitation energy having an r^2 value of 0.93 and a slope of 1.02, and the excited barrier having an r^2 value of 0.85 and a slope of 0.81. Additivity was also tested for the fluorinated **TO** dyes of Chart 1, for which experimental data is available. The results, shown in the Supporting Information, show that the predictions of the additivity model agree well with the experimental results, both for the wavelengths and for the barrier heights (assuming quantum yield in a viscous solution, 90% glycerol, correlates with the barrier height).

The results of this section suggest that additivity is a useful assumption for identifying dye candidates with optimal properties. Additivity has the great advantage of allowing estimates to be made for all possible substitution patterns (nearly 10^6 molecules) using quantum chemical calculations performed on only singly substituted molecules (49 molecules, assuming seven substituents and seven substituent positions).

This prompted us to consider whether it is possible to factor the responses given in Table 3 and 4 into the product of a parameter that captures the strength of the substituent, the **V** of eq 1, and a position dependent proportionality constant, the **A** of eq 2. In section 3.3 we will show that this sort of factorization can not be accurately captured by two parameters.

3.3. Additivity versus Linearity. The above section showed that additivity of substituent effects is a good approximation for the properties of interest here. Because linearity implies additivity, this observation suggests that other consequences of linear response theory may also be present. In particular, it may be possible to write the responses as being proportional to acceptor−donor strength, as captured by some set of Hammett-like parameters. A plot of substituent effect versus the Hammett parameters listed in Supporting Information does not reveal any systematic trends, except for the change in BLA resulting from a substitution at positions P_2, P_3, P_5, and P_6 (see Supporting Information). A more general linear assumption is that of eq 2, which factorizes the response (ΔP) into a product of position dependent constants, **A**, and a set of Hammett-like parameters, **V**. The optimal values of **A** and **V** are obtained from an iterative optimization scheme that minimizes the least-squares difference between the full linear response, ΔP of Tables 3 and 4, and the factorized form of eq 2. Table 5 reports the average percent difference between the original response and the factorized response, with the average being over all substituents and locations. The factorization of the response leads to errors of about 20% for BLA, and much larger errors for the other properties. The smaller error observed for the BLA response is in line with the observation that the BLA is the only property that correlates with existing Hammett parameters. The larger errors resulting from the factorization of eq 2 suggests that, although additivity holds for this system, it is not useful to view the effects of the substituents as being proportional to some measure of acceptor−donor strength. The absence of this potentially useful consequence of linearity is not a hindrance to the design of improved dyes, because additivity alone is sufficient to support computational dye design.

TABLE 5: Average of the Percent Error Calculated for TO Analogues

	% error
BLA[a]	21.3
GSBH[b]	45.9
EE_{planar}[c]	75.1
$EE_{twisted}$[d]	95.3
ESBH[e]	131.5

[a] The bond length alternation. [b] The barrier to rotation in the ground state. [c] Vertical excitation energy at the planar structure (ϕ_1 fixed at 0°). [d] Vertical excitation energy at the twisted structure (ϕ_1 fixed at 90°). [e] Excited-state barrier height.

TABLE 6: Search Results for Dye Candidates with a 30 ± 5 nm Blue Shift Relative to TO

	1[d]	2	3	4	5	6	7
λ[a]	0	0	0	24	209	482	154
BH-10%[b]	0	0	0	1	18	86	34
BH-25%	0	0	0	6	59	154	53
BH-50%	0	0	0	19	144	269	54
<BH[c]	0	0	0	23	194	404	108

[a] Number of molecules that meet the wavelength shift criterion. [b] Number of molecules that are within 10%, 25%, 50% of **TO**'s excited-state barrier height and the targeted wavelength shift. [c] Number of molecules having a excited-state barrier height lower than that of **TO**. [d] Number of substituents on the **TO** framework.

3.4. Design Approach. In the previous section we showed that additivity provides useful estimates for the absorption wavelength and barrier height of the excited-state surface for multiple substitutions on the **TO** framework. In this section, we use additivity to enable an exhaustive search of all possible substitution patterns, using only the responses to single substitutions listed in Tables 3 and 4. Explicit calculations are then performed on the identified dyes to test the additivity approximation. As a design target, we consider the use of substituents to induce spectral shifts while maintaining fluorogenicity. Such spectrally resolved dyes are desirable because they would enable fluorescence imaging to simultaneously monitor multiple biomolecules in a living cell.

Consider, for instance, design of a dye with a 30 nm blue (hypsochromic) shift with respect to the calculated absorption wavelength of unsubstituted **TO**, corresponding to a target of 445 nm. We will assume that retention of the fluorogenic properties is related to the excited-state barrier height. One set of criteria for fluorogenicity that we will consider is to keep the barrier similar to the 0.029 eV of **TO**, with tolerances of 10%, 25% and 50%. This assumes that it is desirable to have a finite barrier such that the molecule does not become overly

TABLE 7: Search Results for the Dye Candidates with a 30 ± 5 nm Red Shift Relative to TO

	1[d]	2	3	4	5	6	7
λ[a]	0	45	851	7837	38416	98036	101904
BH-10%[b]	0	0	24	330	1924	5162	5457
BH-25%	0	1	62	858	4708	12814	13572
Bh-50%	0	6	158	1727	9262	24484	25647
<BH[c]	0	10	326	3454	17977	46144	46688

[a] Number of molecules that meet the wavelength shift criterion. [b] Number of molecules that are within 10%, 25% and 50% of **TO**'s excited-state barrier height and the targeted wavelength shift. [c] Number of molecules having a excited-state barrier height lower than that of **TO**. [d] Number of substituents on the **TO** framework.

TABLE 8: Predicted Maximum Red and Blue Shifts Relative to TO

	no restriction[b]		50%		25%		<BH	
	λ_{red}	λ_{blue}	λ_{red}	λ_{blue}	λ_{red}	λ_{blue}	λ_{red}	λ_{blue}
1[a]	519	468	487	468	487	468	492	469
2	545	461	526	461	501	461	503	461
3	565	454	546	454	536	454	530	454
4	580	447	555	447	550	447	543	447
5	591	442	569	442	564	442	558	442
6	602	442	577	443	573	443	572	444
7	606	445	583	445	583	446	581	449

[a] Number of substituents on the **TO** framework. [b] Maximum change in excited-state barrier height from **TO**.

flexible in the excited state, a feature that could potentially reduce fluorescence even when bound to an aptamer. Another criterion is simply that the barrier be lower than that of TO, such that the molecule rotates to the nonemissive twisted structure in solution. Table 6 lists search results for the dye candidates within ±5 nm of the target absorption wavelength and with various restrictions on the excited-state barrier height. The first candidate that fulfills both criteria (within 10% of the barrier height) has four substituents, three CF_3 groups on the benzothiazole ring at P_1, P_2, and P_4 and a $N(CH_3)_2$ group on the quinoline ring at P_7. The explicitly calculated absorption wavelength and the barrier height are 445 nm and 0.032 eV, respectively, as compared to 450 nm and 0.023 eV obtained assuming additivity. Searching based on the additivity approximation was then able to successfully identify a dye candidate that met the target criteria. A sample list of structures chosen from the 23 dye candidates that have a barrier lower than that of **TO** are shown in Supporting Information.

This suggests that fluorogen design that considers only the substituents needed to induce the desired spectral shift may lead to many dyes that have the target emission wavelength but are not ideal fluorogens.

We next consider design of a dye with a 30 ± 5 nm red (bathochromic) shift with respect to the calculated absorption wavelength of unsubstituted **TO**. Table 7 shows a much larger number of dye candidates compared with Table 6, indicating that red shifts are easier to induce than blue shifts. Table 7 indicates that the number of dyes that satisfy both criteria simultaneously is considerably smaller than those that satisfy only the targeted spectral shift. The selection of a dye from the generated candidates can then be based on criteria such as ease of synthesis. The structures of the 10 dye candidates that have a barrier lower than that of **TO** are shown in Supporting Information.

The additivity model also allows us to address more general questions regarding dye design. For instance, what is the largest red and blue shift that can be obtained for a given multiplicity of substitution? Table 8 lists the predicted maximum red and blue shifts that can be obtained using the **TO** framework and placing various restrictions on the shift in excited-state barrier height. Such analyses allow one to explore the limits of a particular dye framework.

4. Conclusion

The goal of this work is to use computation to aid design of fluorogenic dyes for biological imaging. More specifically, we consider the use of electron acceptor/donor substitutions to alter the dye's photophysical properties. Given the large number of possible substitution patterns on a dye framework, such as that of **TO**, the design would be aided by a better understanding of the structure−property relationships that establish the fluorogenic properties. Previously, we have shown that the fluorogenic properties are related to twisting of the dye, following electronic excitation in solution, from an emissive planar structure to a nonemissive twisted structure. The likelihood that the dye will be a fluorogen is therefore assessed by examining the magnitude of the barrier between the planar and twisted structures.

We began by generating a library of molecules corresponding to single and double substitutions on the **TO** framework. First, we examined whether the calculated properties could be well described with a single bond-length alternation coordinate, BLA, as is true of many properties in linear cyanine dyes.[18] The results indicate that, although the excitation energies correlate fairly well with the BLA, poor correlations are obtained for the excited-state barrier height that establishes the fluorogenic properties. This can be attributed to the excited-state barrier height being more sensitive than BLA to the particular locations of the substituents.

These results prompted us to consider an additivity model that can distinguish between substitutions at various positions. Tests on doubly substituted **TO**'s, and a randomly generated set of triply and quadruply substituted **TO**'s, indicate that additivity is a useful approximation for design purposes. Although additivity suggests an underlying linearity in the substituent effects, additional simplifications stemming from linearity could not be identified. Additivity is sufficient to enable an exhaustive search of substitution patterns. For instance, predictions for the nearly 10^6 possible substitution patterns resulting from placing 7 possible substituents at 7 different substituent locations could be obtained from quantum chemical calculations on only the 49 possible singly substituted molecules. Semiempirical theory is used here because comparison with experimental data suggests it provides more accurate predictions of the substituent effects on the excitation energy than TDDFT. Our past work on nonlinear optical susceptibilities found that additivity within semiempirical theory is a good indicator of additivity in more complex ab initio computations.[20] The design approach used here should then also be applicable within TDDFT and other quantum chemical methodologies, where the computational savings resulting from additivity would be even more essential. Finally, the approach is demonstrated by considering design targets related to using substituents to induce spectral shifts while maintaining an excited-state barrier that is close to that of **TO**, such that the fluorogenic properties may be retained. The practical impact of this is that new dyes can be synthesized having different colors, but by keeping the barrier low, the fluorescence quantum yield will remain correspondingly low. When protein or RNA partners are obtained to bind to these dyes, the net enhancement in fluorescence upon binding will be maximal, leading to high quality fluorescence images and ultrasensitive detection.

Acknowledgment. V.E. and D.J.Y. thank the National Science Foundation (CHE-0719350) for funding. B.A.A. thanks the National Institutes of Health (1 R33 CA9754101-01) and the National Science Foundation (CCF-0330135) for funding. We also thank Sung W. Cho for technical help.

Supporting Information Available: Decomposition of the excited-state barrier into orbital ($E_{HOMO} - E_{LUMO}$), coulomb (J) and exchange (K) energy terms; insensitivity of BLA to substitution position; linear response tables of the ground-state barrier, the excitation energy at $\phi_1 = 90°$, and the BLA coordinate; doubly substituted molecular structures on the same heterocycle with their molecular properties assuming additivity; list of original Hammett parameters and correlation of these parameters with the BLA coordinate; triply and quadruply substituted molecular structures along with their explicitly calculated molecular properties and those assuming additivity; test of additivity for fluorinated **TO** analogues; structures of dye candidates. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Stephens, D. J.; Allan, V. J. *Science (Washington, DC, United States)* **2003**, *300* (5616), 82–86.
(2) Nalbant, P.; Hodgson, L.; Kraynov, V.; Toutchkine, A.; Hahn, K. M. *Science (Washington, DC, United States)* **2004**, *305* (5690), 1615–1619.
(3) Albota, M.; Beljonne, D.; Bredas, J.-L.; Ehrlich, J. E.; Fu, J.-Y.; Heikal, A. A.; Hess, S. E.; Kogej, T.; Levin, M. D.; Marder, S. R.; McCord-Maughon, D.; Perry, J. W.; Rockel, H.; Rumi, M.; Subramaniam, G.; Webb, W. W.; Wu, X.-L.; Xu, C. *Science (Washington, D. C.)* **1998**, *281* (5383), 1653–1656.
(4) Babendure, J. R.; Adams, S. R.; Tsien, R. Y. *J. Am. Chem. Soc.* **2003**, *125* (48), 14716–14717.
(5) Hermann, T.; Patel, D. J. *Science* **2000**, *287* (5454), 820–5.
(6) Li, Q.; Chang, Y.-T. *Nature Protocols* **2006**, *1* (6), 2922–2932.
(7) Brooker, L. G. S.; White, F. L.; Sprague, R. H. *J. Am. Chem. Soc.* **1951**, *73*, 1087–93.
(8) Mujumdar, R. B.; Ernst, L. A.; Mujumdar, S. R.; Lewis, C. J.; Waggoner, A. S. *Bioconjugate Chem.* **1993**, *4* (2), 105–11.
(9) Netzel, T. L.; Nafisi, K.; Zhao, M.; Lenhard, J. R.; Johnson, I. *J. Phys. Chem.* **1995**, *99* (51), 17936–47.
(10) Rye, H. S.; Yue, S.; Wemmer, D. E.; Quesada, M. A.; Haugland, R. P.; Mathies, R. A.; Glazer, A. N. *Nucleic Acids Res.* **1992**, *20* (11), 2803–12.
(11) Armitage, B. A. *Top. Curr. Chem.* **2005**, *253*, 55–76.
(12) Hunt Patricia, A.; Robb Michael, A. *J. Am. Chem. Soc.* **2005**, *127* (15), 5720–6.
(13) Karunakaran, V.; Perez Lustres, J. L.; Zhao, L.; Ernsting, N. P.; Seitz, O. *J. Am. Chem. Soc.* **2006**, *128* (9), 2954–2962.
(14) Silva, G. L.; Ediz, V.; Yaron, D.; Armitage, B. A. *J. Am. Chem. Soc.* **2007**, *129* (17), 5710–5718.
(15) Dietzek, B.; Brueggemann, B.; Pascher, T.; Yartsev, A. *Phys. Rev. Lett.* **2006**, *97* (25), 258301−1–258301/4.
(16) Oudar, J. L.; Chemla, D. S. *Opt. Commun.* **1975**, *13* (2), 164–8.
(17) Chemla, D. S.; Oudar, J. L.; Jerphagnon, J. *Phys. Rev. B: Solid State* **1975**, *12* (10), 4534–46.
(18) Marder, S. R.; Gorman, C. B.; Meyers, F.; Perry, J. W.; Bourhill, G.; Bredas, J.-L.; Pierce, B. M. *Science (Washington, DC, United States)* **1994**, *265* (5172), 632–5.
(19) Yaliraki, S. N.; Silbey, R. J. *J. Phys. Chem. A* **1999**, *103* (14), 2262–2266.
(20) Weibel, J. D.; Yaron, D.; Zyss, J. *J. Chem. Phys.* **2003**, *119* (22), 11847–11863.
(21) Marder, S. R.; Cheng, L. T.; Tiemann, B. G.; Friedli, A. C.; Blanchard-Desce, M.; Perry, J. W.; Skindhoej, J. *Science (Washington, DC, United States)* **1994**, *263* (5146), 511–14.
(22) Runge, E.; Gross, E. K. U. *Phys. Rev. Lett.* **1984**, *52* (12), 997–1000.
(23) Hansch, C.; Leo, A.; Taft, R. W. *Chem. Rev.* **1991**, *91* (2), 165–95.
(24) Dewar, M. J. S.; Jie, C.; Yu, J. *Tetrahedron* **1993**, *49* (23), 5003–38.
(25) AMPAC, 8, 1992−1994 Semichem, Inc. P.O. Box 1649 Shawnee, KS 66222.
(26) Ridley, J.; Zerner, M. *Theor. Chim. Acta* **1973**, *32* (2), 111–34.
(27) Tomlinson, A.; Yaron, D. *J. Comput. Chem.* **2003**, *24* (14), 1782–1788.
(28) Burke, K.; Werschnik, J.; Gross, E. K. U. *J. Chem. Phys.* **2005**, *123*, 062206-1–062206/9.
(29) Crawford, T. D.; Stephens, P. J. *J. Phys. Chem. A* **2008**, *112* (6), 1339–1345.
(30) Marques, M. A. L.; Gross, E. K. U. *Annu. Rev. Phys. Chem.* **2004**, *55*, 427–455.
(31) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A.; Jr.; Vreven, T.; 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 6.0; Gaussian, Inc.: Wallingford, CT, 2004.
(32) Champagne, B.; Guillaume, M.; Zutterman, F. *Chem. Phys. Lett.* **2006**, *425* (1−3), 105–109.
(33) Cherkassky, V.; Mulier, F. *Learning From Data: Concepts, Theory, and Methods*; Wiley-Interscience: New York, 1998.

JP805546S