Chemoinformatic Analysis of a Supertargeted Combinatorial Library of Styryl Molecules

Kerby Shedden,[†] Julie Brumer,[‡] Young Tae Chang,[§] and Gustavo R. Rosania*[,‡]

Department of Statistics, University of Michigan, Ann Arbor, Michigan 48109, Department of Pharmaceutical Sciences, 428 Church Street, University of Michigan College of Pharmacy, Ann Arbor, Michigan 48109-1065, and Department of Chemistry, New York University, Manhattan, New York 10003

Received June 18, 2003

Styryl dyes are fluorescent, lipophilic cations that have been used as specific labeling probes of mitochondria in living cells. For specific applications such as epifluorescence microscopy or flow cytometry, it is often desirable to synthesize fluorescent derivatives with optimized excitation, emission, and localization properties. Here, we present a chemoinformatic strategy suitable for multiparameter analysis of a combinatorial library of styryl molecules supertargeted to mitochondria. The strategy is based on a simple additive model relating the spectral and subcellular localization characteristics of styryl compounds to the two chemical building blocks that are used to synthesize the molecules. Using a cross-validation approach, the additive model predicts with a high degree of confidence the subcellular localization and spectral properties of the styryl product, from numerical scores that are independently associated with the individual building blocks of the molecule. The fit of the data indicates that more complex, nonadditive interactions between the two building blocks play a minor role in determining the molecule's optical or biological properties. Moreover, the observed additive relationship allows mechanistic inferences to be made regarding the structure−property relationship observed for this particular class of molecules. It points to testable, mechanistic hypotheses about how chemical structure, fluorescence, and localization properties are related.

INTRODUCTION

Fluorescent, cell-permeable lipophilic cations are useful for monitoring the structure and function of mitochondria in living cells.[1] Probes of mitochondrial function include well-known fluorescent dyes such as rhodamine1,2,3,[1−4] JC-1,[1,5] as well as cell-permeant, fluorescent dyes of the styryl family.[6−9] The accumulation of lipophilic cations inside the mitochondrial inner matrix has been one of the best-studied, organelle-targeting mechanisms to date.[4,10−12] In the process of oxidative phosphorylation, oxidation of NADH and $FADH_2$ to NAD^+ and FAD is coupled to the pumping of protons across the mitochondrial inner membrane by a series of multienzyme complexes. This pumping mechanism generates a steady-state electrochemical potential across the inner mitochondrial membrane, composed of a pH gradient and a transmembrane voltage. Lipophilic cations accumulate in mitochondria as a function of the transmembrane electrical potential across the mitochondrial inner membrane, in a manner governed by the proton-pumping mechanism and predicted by the Nernst equation.[1,4,5,12,13] Dissipation of the membrane potential is followed by leakage of the probes from the organelle.[1,11,14,15]

Supertargeted chemical libraries refer to a large collection of compounds designed to localize to a specific cellular organelle or subcompartment.[16] The localization of lipophilic cationic probes to mitochondria constitutes one of the best-studied supertargeting mechanisms known to date. Yet, lipophilicity and positive charge are not the only determinants of mitochondrial localization.[17] Nevertheless, the ability to synthesize supertargeted libraries of fluorescent molecules with controlled subcellular localization properties is key to development of novel biosensors for cell biological studies, in vivo imaging, and pharmaceutical screening applications. Understanding the relationship between chemical structure, subcellular distribution, and optical properties is therefore of considerable interest to chemists and biologists.

Recently, we described the facile synthesis of combinatorial libraries of cationic styryl dyes targeting mitochondria as well as other cellular organelles.[17] Different dyes exhibit different subcellular localization properties, and their excitation and emission spectra span a broad range of wavelengths. Styryl dyes normally have a lipophilic pyridinium or quinolinium cation at one end -linked to an aromatic functionality at the other end via a 2 or 4 carbon bridge.[6−9,17] The reported styryl library was synthesized by condensation of 41 aldehydes and 14 pyridinium (2- or 4-methyl) salts. The electron structure of the aromatic systems at both ends of the molecule is conjugated through π -orbitals of bridging carbon−carbon double bonds. Thus, a reasonable starting point for studying the chemical structure−property relationship of styryl molecules involves determining if each building block contributes independently and additively to the physicochemical properties of the resulting product.

To address this question, we applied a simple, statistical regression analysis to study the localization and spectral properties of the styryl library, as encoded by the individual building blocks (A = aldehyde; and B = pyridinium/

* Corresponding author e-mail: grosania@umich.edu.
[†] University of Michigan.
[‡] University of Michigan College of Pharmacy.
[§] New York University.

10.1021/ci0341215 CCC: $25.00
Published on Web 11/06/2003

quinolinium) used to construct the library. The results indicate that nonadditive interaction between A and B moieties across the central double bond has a minimal effect on localization and spectral properties of the styryl molecule. Thus, each individual A or B moiety promotes or inhibits the localization of a particular molecule to mitochondria or contributes to a higher or lower excitation and emission wavelength, by a constant amount, in an additive fashion, and independently from the rest of the molecule. In as much as the Stokes shift[18] dictates that emission is at a higher wavelength than excitation, the numerical contribution of each building block to excitation and emission peaks show some correlation. However, there is little correlation between spectral and localization properties, indicating that development of a toolbox of mitochondria-targeting styryl probes spanning the entire excitation and emission spectrum is possible.

METHODS

Chemistry and Biology. Styryl library synthesis and screening data have been previously published.[17] For staining, UACC-62 cells (obtained from the Developmental Therapeutics Program at the National Cancer Institute) were grown in RPMI medium supplemented with 10% fetal calf serum. For microscopy, cells were plated on 96 well tissue-culture plates (Falcon) overnight, in RPMI plus 10% fetal calf serum, at 37 °C in 5% CO_2/95% air. Cells were incubated with compounds at an approximate concentration of 50 μM for 1 h. Epifluorescence microscopy was performed using an Zeiss Axiovert epifluorescence microscope equipped with a 20X objective, as described.[17] Compounds that yielded no visible intracellular fluorescence were not scored, as lack of fluorescence may be due to low quantum yield, lack of cell permeability, or some unrelated, unknown factor.

Additive Decomposition for Emission and Excitation Spectra. Wavelength values for peak excitation and emission were fit to the additive model $\lambda_{ij} = c + \alpha_i + \beta_j + \epsilon_{ij}$, where ϵ_{ij} denotes an error that is made as small as possible in the fitting process. Using least squares to minimize the function

$$\sum_{ij} = (\lambda_{ij} - c - \alpha_i - \beta_j)^2$$

over all compounds having experimental data yields coefficient estimates α_i for each A group and β_j for each B group. One set of coefficient estimates is obtained for the excitation values, and another set is obtained for the emission values. To predict the wavelength of a new compound formed from A and B groups i^* and j^* the sum $c + \alpha_{i^*} + \beta_{j^*}$ is used, where $c^{ex} = 433$ and $c^{em} = 547$.

Additive Decomposition for Subcellular Localization. Subcellular localization data were converted to binary (0/1) values by assigning a value $G_{ij} = 1$ if compound i,j localized to mitochondria (even if it localized to other compartments as well), and assigning $G_{ij} = 0$ if compound i,j localized exclusively to any nonmitochondrial cellular structure. Compounds with no localization were omitted from this part of the analysis. The α_i and β_j coefficients for A or B groups that are always observed to localize to mitochondria, or that never localize to mitochondria, were set to ± 5, respectively. The binary localization data were analyzed using factorial logistic regression. This method assigns scores α_i and β_j to each A and B group, respectively, so that $\alpha_i + \beta_j > 0$ when compound i,j has a localization value of 1 (i.e. mitochondrial), and $\alpha_i + \beta_j < 0$ when compound i,j has a localization value of 0 (i.e. nonmitochondrial). Specifically, the method maximizes the following function

$$\sum_{ij:Gij=1} \alpha_i + \beta_j - \sum_{ij} \log(1 + \exp(\alpha_i + \beta_j))$$

To predict the localization of a new compound formed from A and B groups i^* and j^* the sum $\alpha_{i^*} + \beta_{j^*}$ is calculated, and the new compound is predicted to be mitochondrial if the sum is positive, and nonmitochondrial if the sum is negative.

Cross-Validation. For both the spectral and localization analysis, cross-validation was used to obtain unbiased estimates of the prediction performance. Each compound was held out in sequence, and the coefficients α_i and β_j were fit to the remaining compounds. These values were then used to form a prediction for the held out compound, then the predicted and experimental values were compared to obtain a measure of the accuracy of prediction. Since the wavelength values are on a continuous scale, the predicted values were compared to the experimental values using Pearson correlation coefficients. The localization values are dichotomous, so the proportion of matching predictions was used to compare predicted and experimental localization values.

Statistical Significance Analysis. To determine the statistical significance of the prediction results, we compared the actual prediction performance to the distribution of performances that would be obtained if the data were randomized. For the localization analysis, performance was measured using the proportion of correctly predicted compounds. The distribution of this proportion when the data are randomized follows the binomial distribution. Thus the *p*-value, which is the likelihood of getting better than the observed prediction results by chance, can be calculated using a table of the binomial distribution. For the spectral analysis, performance was measured using the correlation coefficient between predicted and experimental values. The distribution of these values under randomization can be determined empirically, by repeatedly randomizing the experimental values and repeating the analysis. The proportion of these randomized correlation coefficients that exceed the observed coefficient is reported as the *p*-value.

Similarity Metrics and Cluster Analysis for Data Visualization. The additive decomposition can be used to cluster the data by reordering the rows and columns of the data matrix so that the fitted α_i and β_j coefficients are in increasing or decreasing order. We found this to be the most successful clustering method compared to a number of other clustering algorithms that we tested empirically on the styryl data. The relationship between different A and B functionalities was calculated using a variety of commonly used similarity metrics (between groups, within groups, nearest neighbor, furthest neighbor, etc.). The resulting relationships were then organized into categories using a variety of hierarchical clustering algorithms. None of the similarity metrics and hierarchical clustering algorithms tested yielded results that were as satisfactory as those obtained with the additive decomposition analysis, for reasons explained in the text.

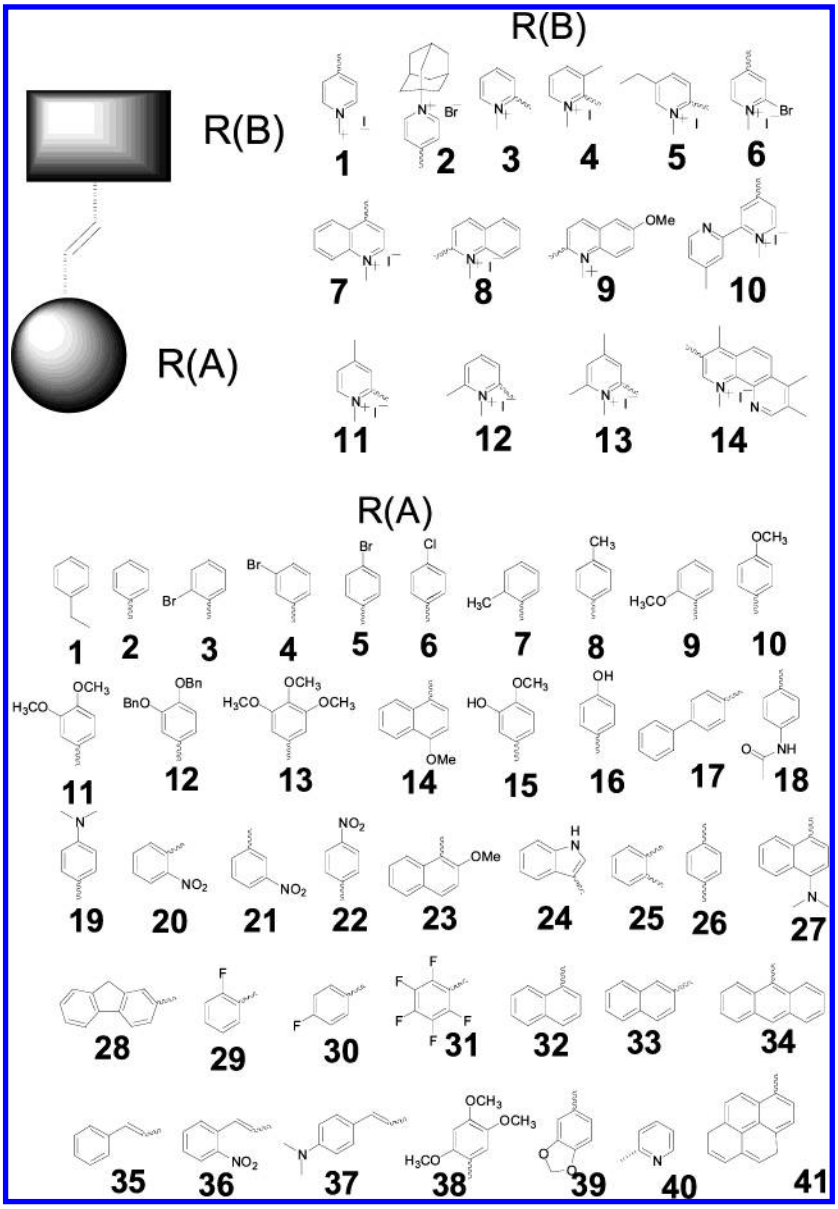

Figure 1. The structure of the styryl library, comprised of all possible pairwise combinations of A and B groups.

RESULTS

Analysis of Fluorescence Excitation and Emission. The chemical structure of the styryl library[17] is illustrated in Figure 1. Initial analysis focused on measurements of peak emission and excitation wavelength obtained for all styryl products showing a single, localized peak (there were 256 such compounds for emission wavelength and 193 compounds for excitation wavelength). Peak emission and excitation wavelengths were found to vary over almost the entire visible range. We will denote the wavelength for the styryl compound formed by joining A group i with B group j as λ_{ij} (or more specifically, λ_{ij}^{ex} for excitation wavelength and λ_{ij}^{em} for emission wavelength). The additive model $\lambda_{ij} = c + \alpha_i + \beta_j + \epsilon_{ij}$ was fit to the data using least squares (see methods), yielding parameters α_i^{ex}, α_i^{em}, β_j^{ex}, and β_j^{em} that quantify the influence of each A and B group on the spectrum of the styryl product. The resulting fitted values $\lambda_{ij}^{ex} = c^{ex} + \alpha_i^{ex} + \beta_j^{ex} + \epsilon_{ij}^{ex}$ and $\lambda_{ij}^{em} = c^{em} + \alpha_i^{em} + \beta_j^{em} + \epsilon_{ij}^{em}$ showed good correlation with the true values (Figure 2). Using cross-validation, each compound was held out in sequence, the model was trained using the remaining compounds, and then the wavelength value for the held-out compound was predicted based on the resulting fitted model. This process produced correlations between measured and predicted values of $\rho^{em} = 0.78$ (emission) and $\rho^{ex} = 0.69$ (excitation).

The degree of correlation between predicted and experimental peak wavelengths was highly statistically significant for both excitation and emission values. Randomizing the compounds 1000 times (see methods) yielded a null distribution of correlation coefficients with 95th percentile 0.12 and maximum value 0.23—far smaller than the observed values of both spectra (0.78 and 0.69) given above. We also sought to determine whether the spectral properties of the styryl product vary according to the identity of both the A and B group, or whether only one of the two groups has a differential influence. To assess this, we refit the model described above while holding either $\alpha_{ij} = 0$ (allowing no differential effect of the A group on peak wavelength) or $\beta_{ij} = 0$ (allowing no differential effect of the B group on peak

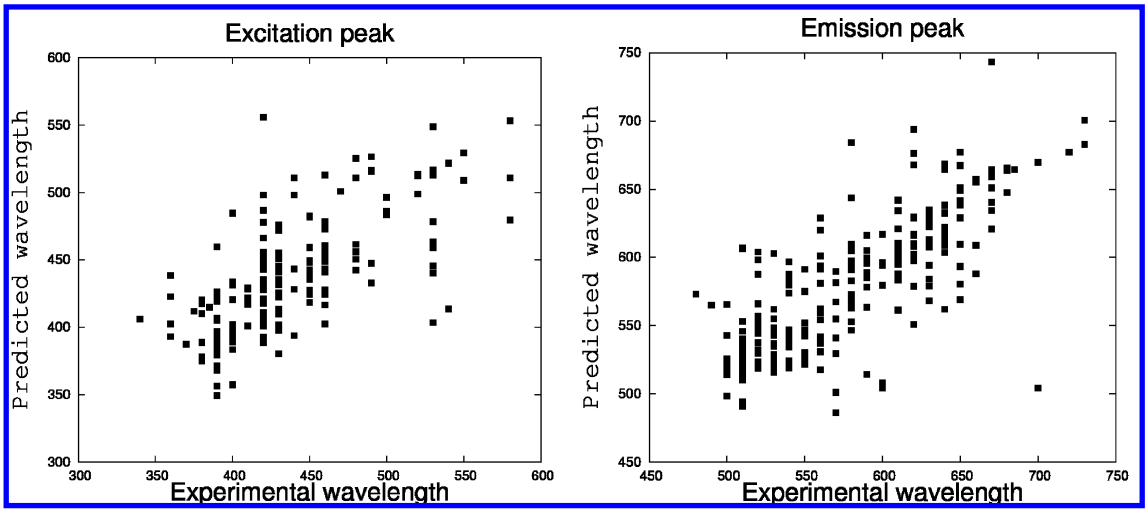

Figure 2. Predicted versus experimentally determined values for peak emission (left) and excitation (right) wavelengths in the styryl library. The predictions were made ignoring interactions between the two functional moieties in the styryl compound. The predicted values were obtained without bias, by holding out the data point to be predicted when training the model.

wavelength). The resulting fitted values showed much lower correlation with the true values, compared to the additive model that allows differential effects for both groups. For emission wavelengths, the correlation between predicted and observed values based on the identities of both the A and B group were 22% higher than the correlation based only on the A group identity and were 95% higher than the correlation based only on the B group identity. For excitation wavelengths, the corresponding values are 73% and 53%.

Based on this analysis, the contribution of the A and B functional groups to the fluorescence of the styryl product can be quantified using the fitted coefficients α_i^{ex}, β_j^{ex}, α_i^{em}, and β_j^{em}. Since these coefficients are on a scale without origin, we use the first A group and B group as a baseline, so $\alpha 1^{ex} =$, $\beta 1^{ex} = 0$, and so on. Table 1 contains the model coefficients of all A and B groups for peak excitation and emission wavelength. Positive coefficients indicate that the corresponding A or B group reddens the peak wavelength, with a greater magnitude indicating a greater degree of reddening. The range for the A group coefficients is around 150 nm for both excitation and emission peak wavelength, which means that by changing the identity of the A group in the styryl compound, one can systematically shift the peak wavelength by around half the width of the visible spectrum. For example, changing the A group from 37 to 12 is associated with roughly a 140 nm shift in peak emission wavelength, across a diverse range of B groups. Although shifts greater than 140 nm may be seen in specific pairs of compounds, the 140 nm shift is notable in that it is seen consistently in a diversity of compounds where only the A group varies. Changes in the B group also lead to sizable, though smaller changes. For example, changing the B group from 8 to 5 is associated with roughly a 70 nm shift across a diverse range of A groups. By changing both the A and B groups at the same time, the whole visible spectrum is covered.

Analysis of Complex Spectra. One possible application of the additive model for peak wavelengths is to make inferences about styryl products whose spectral properties deviate from the norm. These styryl products may represent failed synthesis, reactions that yield multiple fluorescent products, or formation of dye aggregates with complex optical properties. More interestingly, they may also represent products with conformation-dependent or environment-sensitive optical properties that could be exploited for biosensing applications. For the initial fitting and testing of the model, spectra of compounds exhibiting multiple peaks

Table 1: Influence of A and B Groups on Peak Excitation and Emission Wavelengths, and on Subcellular Localization, Inferred from Measurements on Styryl Molecules[a]

A Groups

	α_i^{ex}	α_i^{em}	α_i^{mito}		α_i^{ex}	α_i^{em}	α_i^{mito}
1	0.0	0.0	−0.2	22	−52.8	−10.3	5.0
2			5.0	23	6.1	−10.3	−5.0
3	35.7	46.3	5.0	24	−7.2	−24.2	−1.7
4	−34.0	23.5	−5	25	−5.3	−21.9	−5.0
5	−32.3	26.7	5.0	26	13.2	24.7	5.0
6	4.9	−9.5	5.0	27	11.1	72.0	−0.3
7	−44.1	38.5	5.0	28	2.4	44.9	−5.0
8	10.9	38.5	5.0	29		−8.0	5.0
9	−36.7	−17.6	5.0	30	7.4	50.1	5.0
10	8.1	−23.7	1.3	31	−22.4	6.9	5.0
11	−14.3	−26.7	5.0	32	−36.1	−15.0	5.0
12	−36.7	−22.7	−3.3	33		−26.5	5.0
13	−44.2	0.2	5.0	34	14.1	58.8	−5.0
14	−1.6	0.3	−1.2	35	0.0	−8.3	−1.8
15	−46.9	−33.1	−0.6	36	−25.3	3.8	5.0
16	−36.3	−24.8	−1.9	37	82.9	122.1	−0.9
17	−14.3	−46.6	−5.0	38	−15.6	22.2	−1.3
18	−24.3	51.2	0.8	39	−28.3	−25.3	0.0
19	26.0	49.6	−0.2	40	−18.1	64.8	
20	6.5	44.5	−2.7	41	26.1	46.9	
21	−68.0	21.6	0.8				

B Groups

	β_j^{ex}	β_j^{em}	β_j^{mito}		β_j^{ex}	β_j^{em}	β_j^{mito}
1	0.0	0.0	1.9	8	53.0	64.8	0.4
2	4.9	4.3	2.0	9	−0.1	35.0	−2.0
3	−12.7	−2.7	4.2	10	1.6	19.8	0.8
4	−32.3	−23.2	3.0	11	−4.2	−4.1	2.1
5	−1.6	−7.6	2.2	12	−10.2	−1.7	2.7
6	27.7	3.5	−5.0	13	−1.1	−1.8	2.5
7	29.6	63.7	3.5	14	100.6	57.9	0.0

[a] Greater values of α_i^{ex}, α_i^{em}, β_j^{ex}, and β_j^{em} indicate greater peak wavelength. Greater values of α_i^{mito} and β_j^{mito} indicate stronger mitochondrial localization. A groups 40 and 41 were not screened for localization. A groups 2, 29, and 33 only formed fluorescent products with a single B group, so they were not included in the spectral analysis.

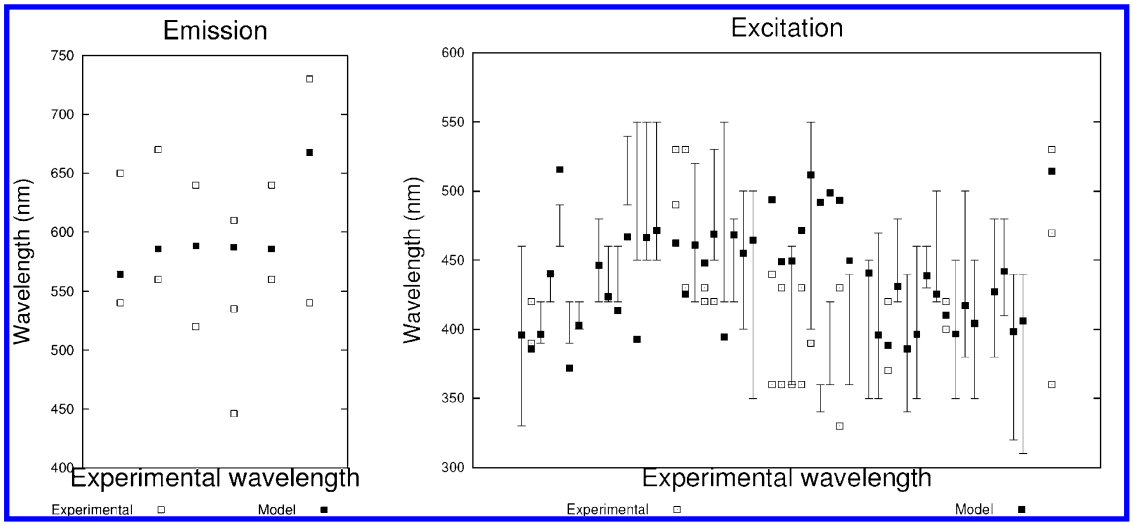

Figure 3. Experimental and predicted peak emission and excitation wavelengths for compounds with complex spectra. Each vertical band corresponds to a single such compound, empty squares represent measured excitation or emission peak values, and filled squares indicate the predicted peak wavelength according to the additive model. Multiple localized peaks in the experimentally determined spectra are shown as multiple unfilled quares, and a single broad peak is shown as a vertical error bar.

or poorly defined peaks were ignored. Thus, using the model fit to the compounds with simple spectra, the peak wavelength can be predicted for compounds with complex spectra and compared to the measured spectra. Figure 3 shows the predicted peak emission and excitation wavelengths for the compounds with complex spectra, along with the experimentally determined peak wavelengths (Each vertical band represents a single compound; the experimental data are shown as either a vertical error bar for a poorly defined broad peak or as multiple empty squares for several localized peaks.). For the 38 products with broad peaks, it is seen that 29/38 of the predicted values fall somewhere within the peak, suggesting that in most cases, products with complex spectra also follow the additive relationship observed for products with single excitation or emission peaks.

We found this result to be statistically significant at $p < 0.0001$. For simulation, we established the probability that the predicted excitation/emission values would fall within the measured, complex values base on pure chance. For this purpose, the predicted values were randomly assigned to the 38 complex spectra. The number of times that the predicted and measured spectra would overlap was scored, and the entire procedure was iterated $>10^4$ times. Based on this simulation, for 38 broad peaks randomly assigned to the corresponding 38 predicted values, on average only 19 of the predicted values are covered (with 95th percentile point 23 and maximum value of 28 in 10^4 random assignments). In addition, for a number of the compounds with multiple peaks (e.g. the leftmost two in the emission data), the predicted peak is much closer to one of the experimental peaks compared to the others, suggesting that the complex spectra may be due to the presence of multiple fluorescent products and that at least one of these products corresponds to the expected product.

Analysis of Mitochondrial Localization. Localization analysis was focused on how A and B groups determine nonmitochondrial from mitochondrial localization (indiscriminately of whether mitochondrial localization is specific), by calculating the proportion of all compounds that are correctly predicted as localizing to mitochondrial or nonmitochondrial structures. As was the case for the compounds' fluorescence properties, this analysis aimed to determine whether A and B groups determined localization in an independent, additive fashion, with each A and B group contributing toward localization by a constant amount and independent of the rest of the molecule. Measurements of subcellular localization were made for 147 of the styryl compounds, as previously described. Due to the cationic nature of the B groups, many of the styryl compounds are expected to accumulate in mitochondria. While this is true of roughly half of the compounds, many compounds localize to nucleus, nucleolus, cytosol, ER, and to cytoplasmic granules.[17]

Unlike measurement of fluorescence excitation or emission peaks, localization to mitochondria is determined by visual inspection and was scored in a binary fashion[17] (Reaction products localizing to mitochondria are given a value of 1, while those that do not are given a value of 0.). To analyze these data, we used a factorial logistic regression approach (see methods) to establish if A and B groups additively and independently contribute to localization. Briefly, this technique assigns quantitative scores to each A group (α_i^{mito}) and to each B group (β_j^{mito}), in such a way that $\alpha_i^{\text{mito}} + \beta_j^{\text{mito}}$ is positive for compounds with mitochondrial localization and is negative for compounds lacking mitochondrial localization. Good prediction performance would suggest that indeed A and B groups contribute to localization in an additive, independent fashion. While in principle one could apply the same analysis to organelles other than mitochondria, there are not enough localizations to specific nonmitochondrial organelles in the styryl library to calculate reliable statistics (data not shown), so the present analysis focuses only on mitochondria.

The predictive performance of this method was assessed using cross-validation, by calculating the proportion of all compounds that are correctly predicted. For cross-validation, each styryl product was set aside in sequence, and the factorial logistic model was fit to the remaining data. Then the resulting A and B scores for the held-out compound were summed. If this sum was positive, the held-out compound was predicted to be mitochondrial, while if the sum was negative, the held-out compound was predicted to be

Table 2: Raw Localization Data, Estimated Affinity Coefficients, and Results of a Prediction Analysis[a]

	B_j	3	7	4	12	13	5	11	2	1	10	8	14	9	6
	b_j	4.2	3.5	3.0	2.7	2.5	2.3	2.1	2.0	1.9	0.8	0.4	0.1	-2.0	-5.0
A_i	a_i														
3	5.0	9.2	8.5	8.0	7.7	7.5	7.3	7.1	7.0	6.9	5.8m	5.4m	5.1	3.0m	0.0
7	5.0	9.2	8.5m	8.0	7.7	7.5	7.3	7.1	7.0	6.9	5.8	5.4	5.1	3.0	0.0
8	5.0	9.2	8.5m	8.0	7.7	7.5	7.3	7.1	7.0	6.9	5.8	5.4	5.1	3.0	0.0
9	5.0	9.2	8.5m	8.0	7.7	7.5	7.3	7.1m	7.0	6.9	5.8m	5.4m	5.1	3.0	0.0
11	5.0	9.2	8.5	8.0	7.7	7.5	7.3	7.1	7.0m	6.9	5.8	5.4	5.1	3.0	0.0
22	5.0	9.2	8.5	8.0	7.7	7.5	7.3	7.1	7.0m	6.9	5.8	5.4	5.1	3.0	0.0
30	5.0	9.2	8.5	8.0	7.7	7.5	7.3	7.1	7.0	6.9	5.8	5.4	5.1	3.0m	0.0
31	5.0	9.2	8.5m	8.0	7.7	7.5	7.3	7.1	7.0m	6.9	5.8m	5.4m	5.1	3.0m	0.0
32	5.0	9.2	8.5m	8.0	7.7m	7.5m	7.3	7.1m	7.0m	6.9m	5.8m	5.4m	5.1	3.0m	0.0
33	5.0	9.2	8.5m	8.0	7.7m	7.5m	7.3	7.1m	7.0	6.9m	5.8m	5.4m	5.1	3.0m	0.0
36	5.0	9.2	8.5	8.0	7.7	7.5	7.3	7.1	7.0m	6.9	5.8	5.4	5.1	3.0	0.0
10	1.3	5.4	4.8	4.2	4.0m	3.8	3.5	3.4m	3.3	3.2	2.1m	1.7m	1.3	-0.8o	-3.7
18	0.8	5.0	4.3m	3.8	3.6	3.4	3.1	2.9	2.9	2.7	1.6	1.3m	0.9	-1.2o	-4.2
21	0.8	5.0	4.3m	3.8	3.6	3.4	3.1	2.9	2.9	2.7	1.6	1.3m	0.9	-1.2o	-4.2
39	0.0	4.2	3.5	3.0	2.7	2.5m	2.3	2.1m	2.0m	1.9	0.8m	0.4o	0.1	-2.0	-5.0
19	-0.2	4.0m	3.3m	2.7	2.5m	2.3m	2.0m	1.9m	1.8m	1.7m	0.6m	0.2o	-0.1o	-2.2o	-5.2
1	-0.2	4.0	3.3	2.7	2.5	2.3	2.0	1.9	1.8o	1.7o	0.6m	0.2m	-0.1	-2.2m	-5.2
27	-0.3	3.9	3.2m	2.6	2.4	2.2m	1.9	1.8m	1.7m	1.6o	0.5	0.1m	-0.3	-2.4o	-5.3
15	-0.6	3.6	2.9	2.3	2.1	1.9	1.6	1.5	1.4	1.3	0.2o	-0.2m	-0.5	-2.7	-5.6
37	-0.9	3.2	2.6	2.0m	1.8	1.6	1.3m	1.2o	1.1m	1.0	-0.1o	-0.5	-0.9m	-3.0o	-5.9
14	-1.2	3.0	2.3m	1.8	1.6m	1.4m	1.1	0.9m	0.9o	0.7m	-0.3	-0.7o	-1.1	-3.2o	-6.2
38	-1.3	2.9	2.2o	1.7	1.5	1.2	1.0	0.8	0.7m	0.6	-0.5m	-0.8	-1.2	-3.3o	-6.3
24	-1.7	2.5	1.8m	1.3	1.1	0.8	0.6	0.4m	0.3m	0.2o	-0.9	-1.2o	-1.6o	-3.7	-6.7o
35	-1.8	2.4	1.7	1.1	0.9m	0.7	0.4	0.3	0.2	0.1	-1.0o	-1.4	-1.7	-3.9o	-6.8
16	-1.9	2.3	1.6	1.1	0.8	0.6	0.4	0.2	0.1m	0.0	-1.1o	-1.4o	-1.8	-3.9o	-6.9
20	-2.7	1.5	0.8m	0.3	0.0	-0.2	-0.4	-0.6	-0.7	-0.8	-1.9o	-2.2o	-2.6o	-4.7o	-7.7
12	-3.3	0.9m	0.2m	-0.3	-0.6o	-0.8o	-1.0o	-1.2o	-1.3o	-1.4m	-2.5o	-2.9o	-3.2	-5.3o	-8.3
23	-5.0	-0.8o	-1.5o	-2.0o	-2.3o	-2.5o	-2.7o	-2.9o	-3.0o	-3.1m	-4.2	-4.6o	-4.9	-7.0	-10.0
4	-5.0	-0.8	-1.5	-2.0	-2.3	-2.5	-2.7	-2.9	-3.0	-3.1	-4.2	-4.6	-4.9	-7.0o	-10.0
17	-5.0	-0.8	-1.5	-2.0	-2.3o	-2.5o	-2.7	-2.9o	-3.0	-3.1	-4.2	-4.6o	-4.9	-7.0o	-10.0
25	-5.0	-0.8	-1.5	-2.0	-2.3	-2.5	-2.7	-2.9	-3.0o	-3.1	-4.2	-4.6	-4.9	-7.0	-10.0
28	-5.0	-0.8	-1.5	-2.0	-2.3	-2.5	-2.7	-2.9	-3.0	-3.1	-4.2	-4.6o	-4.9	-7.0o	-10.0

[a] The first column and first row contain the A group and B group labels. The second column and second row contain the estimated A group and B group affinity coefficients (a_i and b_j, respectively). The interior of the table contains the value $s_{ij} = a_i + b_j$ for each compound (where positive values of s_{ij} indicate predicted localization to mitochondria, and negative values of s_{ij} indicate predicted localization to a nonmitochondrial compartment). The subscript m indicates experimentally determined mitochondrial localization. The subscript o indicates experimentally determined nonmitochondrial localization. The darkened boxes indicate correctly predicted mitochondrial or nonmitochondrial localization under cross-validation.

nonmitochondrial. Table 1 gives the fitted model coefficients α_i^{mito} and β_j^{mito} for mitochondrial localization, and Table 2 gives the values for $\alpha_i^{\text{mito}} + \beta_j^{\text{mito}}$ in relation to the empirically determined mitochondrial localization. Positive values of these coefficients suggest that the corresponding A or B group confers mitochondrial localization to compounds of which it is a part (The numbers +5 or −5 were used for groups that conferred mitochondrial localization in every case, or no case, respectively.). The range in these coefficients is around 4.6 (excluding values fixed at +5 or −5), indicating that by changing the identity of the A group, an odds ratio of around 100 can result (The odds ratio is the probability ratio of mitochondrial localization to nonmitochondrial localization.). The baseline performance of the method was 104 correct out of 145, or 72%. This number is highly statistically significant compared to random guessing (p-value $\sim 10^{-7}$). Thus, across the entire library, A and B moieties appear to contribute to localization in an independent, additive fashion.

Since statistical power for assessing interactivity increases when a greater number of combinations is observed, we also considered error rates for subsets of A and B groups where localization could be determined for a minimum of number of products (represented by the coefficient k). The percentage of correct predictions increases from 72% ($k = 0$; comprising the entire data set) to 88% ($k = 10$; comprising those A and B groups that yielded the greatest number of localizable products; Table 2). This suggests that to a high degree, mitochondrial localization is determined by independent contributions from the A and B functional groups. The relatively higher overall error rate (28% for $k = 0$), compared to the error rate for the subset of compounds comprised of groups observed in many distinct configurations (12% for $k = 10$), can be attributed to training error in the coefficients α_i^{mito} and β_j^{mito}, which is reduced as k increases.

The differential influence of both A and B groups was also calculated for the localization properties, by a similar method used to determine the differential influence of A and B groups to spectral properties (as discussed in the previous section). Unlike excitation or emission peaks, differential localization appears to be influenced mostly by contributions from the A group. Table 3 shows the prediction performance for mitochondrial localization based on the identity of the A group alone (columns 4 and 5) and based on the B group alone (6 and 7). The latter prediction is not significantly better than chance, while the former is nearly comparable to prediction based on both groups. This suggests that localization varies consistently with the identity of the A group, while the B groups are more or less interchangeable. One explanation for this may be that the positive charge in the B moiety draws the compound toward mitochondria (equally for all 14 B groups), while certain of the A groups are drawn toward other organelles or otherwise prevent accumulation of the molecule to mitochondria. Thus, for this group of compounds, the A group ultimately determines whether ac-

Table 3: Prediction Performance based on Cross-Validation for Mitochondrial Localization in the Styryl Library, based on Factorial Logistic Regression[a]

	A&B		A only		B only	
k	#correct/ #total	prop. correct	#correct/ #total	prop. correct	#correct/ #total	prop. correct
0	104/145	0.72	95/145	0.66	82/145	0.57
2	95/136	0.70	87/136	0.64	75/136	0.55
4	86/115	0.75	79/115	0.69	67/115	0.58
6	52/66	0.79	47/66	0.71	41/66	0.62
8	43/50	0.86	42/50	0.84	31/50	0.62
10	14/16	0.88	13/16	0.81	7/16	0.44

[a] Predictions were based on both the A and B group (columns 2 and 3), the A group only (columns 4 and 5), or the B group only (columns 6 and 7). Rates of correct prediction are given for the set of compounds in which the A and B group both belong to at least *k* compounds having localization data, for various values of *k*.

cumulation is mostly in the mitochondrial direction or mostly in the direction of the other, nonmitochondrial organelle.

Data Clustering and Visualization. For visualization, the analysis presented here provides a logical, intuitive way of clustering the data. Figure 4 shows the clustered, experimental wavelengths for peak emission and excitation, respectively, while Figure 5 shows the clustered localizations, as determined empirically and based on the sorted α_i^{mito} and β_j^{mito}. To generate the tables, the additive decomposition analysis is applied, and rows and columns of the data matrix are sorted so that the alpha coefficients increase from bottom to top, and the beta coefficients increase from left to right. According to the additive model, compounds formed from the A and B group having the largest α_i^{ex} and β_j^{ex} (or α_i^{em} and β_j^{em}) will have the greatest wavelength at the top right corner of the matrix, and the wavelength will decrease as either α_i or β_j decreases. Thus the color will shift from red to blue while moving vertically from top to bottom or horizontally from right to left in the reordered table. The color will move more rapidly from red to blue while moving along the diagonal from the top right to the lower left of the table. The rate at which the color varies can easily be ascertained by inspecting the reordered table. A continuous rate of change indicates that all colors are roughly equally represented, while skew or sudden changes indicate that certain bands of the spectrum predominate, and others are under-represented.

As with the spectral data, the localization data can also be clustered and visualized by the additive decomposition method (Figure 5). To generate the localization table, the additive decomposition analysis is applied, and rows and columns of the data matrix are sorted so that the alpha coefficients increase from top to bottom, and the beta coefficients increase from left to right. According to the additive model, compounds formed from the A and B group having the largest α_i^{mito} and β_j^{mito} will have the greatest probability of being localized to mitochondria at the bottom right corner of the matrix, and the probability of finding a mitochondrial-localized product will decrease as α_i^{mito} and β_j^{mito} decrease. Thus the localizations shift from O to M while moving vertically from top to bottom or horizontally from left to right in the reordered table. For the clustered localization, the differential influence of A and B groups on mitochondrial vs nonmitochondrial localization can be readily visualized (Figure 5). It is evident that for every B, there are both mitochondrial (M) and nonmitochondrial (O) styryl products. This is consistent with the B group exerting a minimal differential influence on localization. Conversely, for the A groups, three different clusters can be observed: Clusters 1, 2, and 3 correspond to A groups exclusively associated with M, M/O, or with O, in the respective order.

We compared the results based on the additive decomposition to results of more conventional clustering methods, including two-way hierarchical clustering and a Monte Carlo search procedure that maximizes the local similarity within a neighborhood of wavelengths (data not shown). Since we have already established that the additive model fits the data reasonably well, it is not surprising that the additive decomposition produced clustering results surpassing the other methods that we tried, at least from a subjective, visualization viewpoint (the other methods produced rearrangements with several isolated clusters of high or low frequency compounds, rather than the global gradient produced by the additive model). The reason for this may be that other clustering algorithms (e.g. most implementations of hierarchical or agglomerative clustering) are greedy, meaning that they make changes to the arrangement in a sequence of small steps, with each step influenced only by local features of the cluster quality. The additive model, on the other hand, is by nature a global method, since all coefficients are sensitive to changes in any other coefficient, through the least-squares fitting process.

Another advantage of using the additive decomposition for clustering is that it provides a unique solution, with fixed reference points—the upper right corner is always the reddest part of the table, and the lower left corner is always the bluest part of the table. On the other hand, other methods do not provide a unique solution, and there are many transformations, such as a vertical or horizontal flip, that return a distinct, but equally valid solution. Yet another advantage is that the additive method easily handles missing information (i.e. compounds lacking experimental data), since it only is necessary to observe a limited number of compounds to identify and estimate the additive coefficients. As a caveat, the ability to use the additive decomposition to effectively cluster and visualize the data simply reflects the goodness of the additive fit that has already been found to characterize the data set under study. If the effect of chemical groups A and B on the wavelength and localization of the styryl molecules were not additive, it would be impossible to reorder the rows and columns so that a gradient is obtained. Nevertheless, for analysis of the styryl compounds, clustering by additive decomposition clearly yields the best visualization results.

Analysis of Multiparameter Labeling. Multiparameter labeling refers to the ability to use different fluorescent probes to simultaneously monitor different cellular organelles in a single living cell. For multiparameter labeling, it is important to have probes that localize to different organelles, but also their fluorescence excitation and emission spectra need to be nonoverlapping. In other words, the optical properties of the probes should allow discriminating one probe from the other within a single living cell or cell population, using optical filters. To determine whether a combinatorial library of styryl molecules may serve as a toolbox suitable for multiparameter labeling, it is also

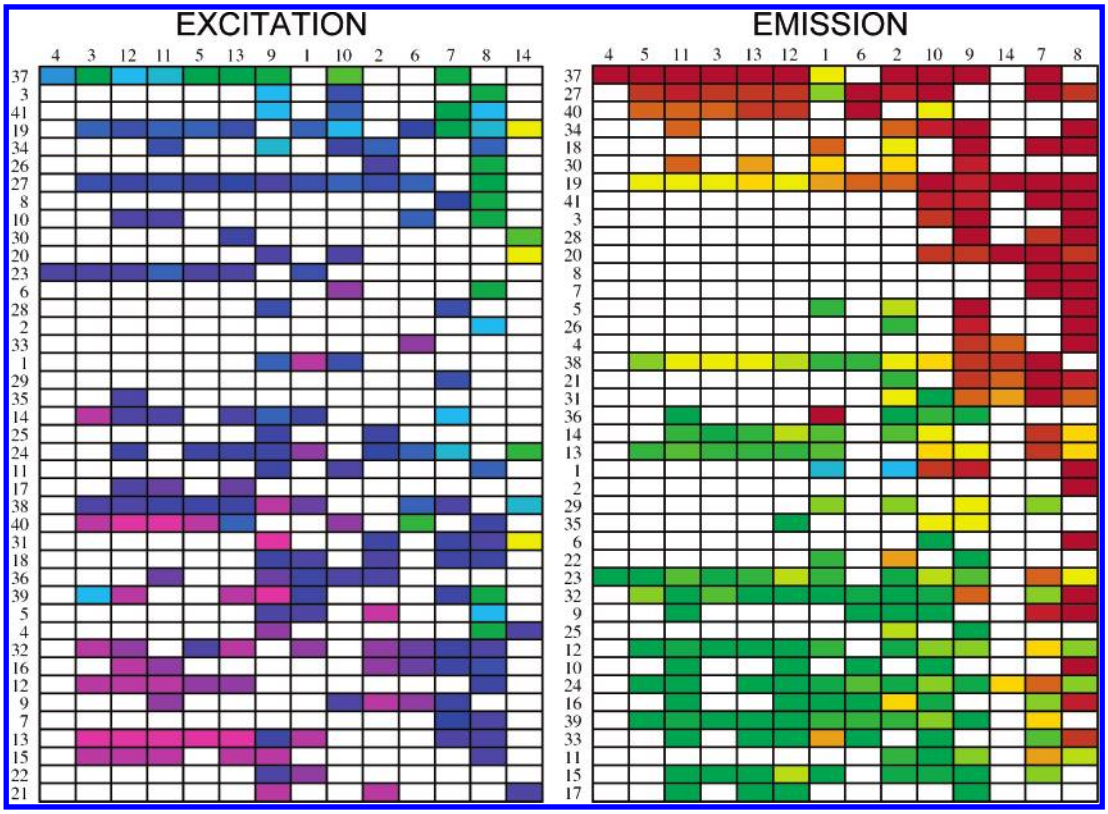

Figure 4. Clustered peak wavelength excitation (left) and emission (right) plots. Each wavelength was color-coded with the corresponding color, so that trends could be easily visualized.

	B6	B9	B14	B8	B10	B1	B2	B11	B5	B13	B12	B4	B7	B3	
A4		O													
A17		O		O				O		O	O				
A23				O		M	O	O	O	O	O	O	O	O	
A25							O								
A28		O		O											
A12		O		O	O	M	O	O	O	O	O		M	M	3
A20		O	O	O	O								M		
A16		O		O	O		M								
A35		O			O						M				
A24	O		O	O		O	M	M					M		
A38		O			M		M						O		
A14		O		O		M	O	M		M	M		M		
A37		O	M		O		M	O	M			M			
A15				M	O										
A27		O		M		O	M	M		M			M		2
A1		M		M	M	O	O								
A19		O	O	O	M	M	M	M	M	M	M		M	M	
A39				O	M		M	M		M					
A18		O		M									M		
A21		O		M									M		
A10		O		M	M			M			M				
A3		M		M	M										
A7													M		
A8													M		
A9				M	M			M					M		
A11							M								
A22							M								1
A30		M													
A31		M		M	M		M						M		
A32		M		M	M	M	M	M		M	M		M		
A33		M		M	M	M		M		M	M		M		
A36							M								

Figure 5. Clustered mitochondrial (M) and nonmitochondrial (O) localizations. Three groups are indicated, highlighting relative differences in mitochondrial affinity: group 1 is predominantly mitochondrial; 2 is both mitochondrial and nonmitochondrial; 3 is predominantly nonmitochondrial.

important to determine that—for each localization (mitochondrial vs nonmitochondrial)—a broad range of fluorescence excitation and emission wavelengths are available.

For this purpose, a joint analysis of fluorescence and localization properties was carried out. Initially, a bivariate,

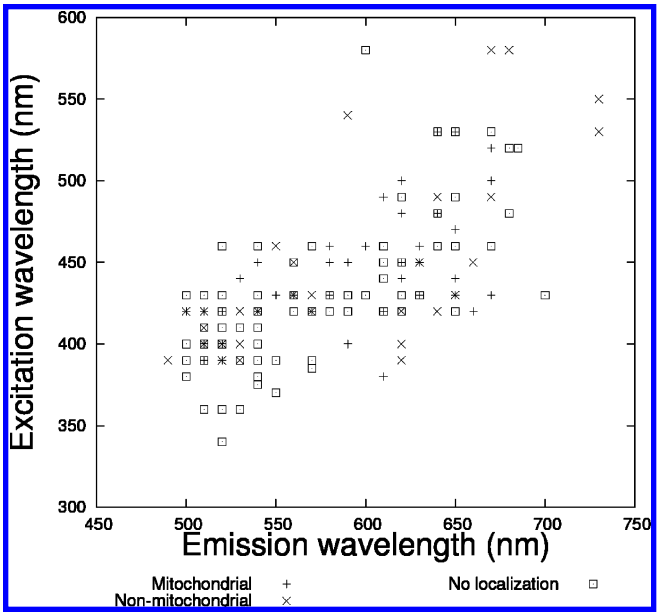

Figure 6. Bivariate plot of excitation and emission peak wavelength distribution of styryl products, indicating different localizations.

excitation vs emission plot was used to compare the fluorescence properties of mitochondrial and nonmitochondrial compounds, together with the fluorescence properties of compounds that do not localize to any cellular compartments (Figure 6). Styryl products that show mitochondrial localization excite anywhere between 380 and 540 nm and emit anywhere between 500 and 660 nm. Products that do not show mitochondrial localization or that do not localize altogether excite from 340 to 580 nm and emit anywhere from 480 to 730. Thus mitochondrial-targeting styryl dyes

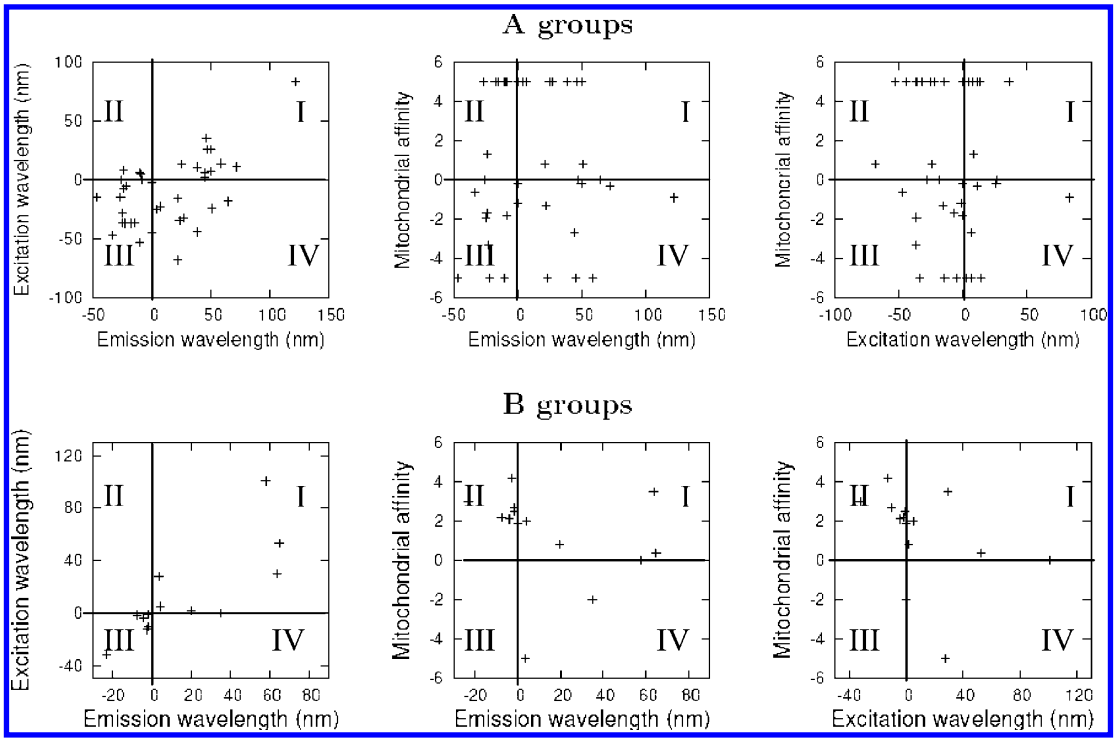

Figure 7. Bivariate plots of excitation/emission (left panels), mitochondrial affinity/emission (middle panels), and mitochondrial affinity/excitation (right panels) for the individual A (top) and B (bottom) groups. For clarity, each quadrant in the plot is indicated with roman numerals.

show quite a broad spectral, but the nonmitochondrial targeting dyes show an even broader range.

To determine the contribution of A and B moieties to this trend, we analyzed bivariate plots of A and B moieties looking for correlations between excitation-localization or emission-localization contributions (Figure 7). As a positive control, we started by analyzing the excitation−emission contribution, which should show correlation based on the Stokes shift (according to quantum mechanical principles, a molecules emission wavelength must be higher than the excitation wavelength[18]). Accordingly, the plot reveals most of the data points lying in quadrants I, III and IV, indicating that A or B moieties that are red shifted in the excitation are unlikely to be blue shifted in the emission. Referring to the localization/emission plots for the A groups, one can see an equal distribution of data points on quadrants I, II, III, and IV indicating that localization and fluorescence contributions are not correlated. On the other hand, for the B groups, data points fall on quadrants I, II, and IV but not on III. While this may suggest a certain correlation between fluorescence and mitochondrial contributions for the B groups, B groups do not exert a differential influence on mitochondrial localization indicating that this result is not statistically significant.

For microscopy and live cell applications, dyes that excite and fluoresce at 480 nm or higher are most desirable, as intracellular NADH and FADH leads to high background autofluorescence at lower wavelengths. Thus, by virtue of its mitochondrial localization/fluorescence properties, styryl libraries appear to be naturally biased toward finding good fluorescent reporters for mitochondrial visualization in the visible wavelengths. But for the most part, there is no strong association between the model coefficients for peak wavelength (either emission or excitation) and mitochondrial localization. Therefore the functional A and B groups used to build the styryl library appear to independently confer shifts in spectral and subcellular localization properties. For example, among the A groups, group 17 confers mitochondrial repulsion and a bluer λ_{ij}^{em}; group 34 confers mitochondrial repulsion and a redder λ_{ij}^{em}; group 33 confers mitochondrial attraction and a bluer λ_{ij}^{em}; and group 30 confers mitochondrial attraction and a redder λ_{ij}^{em}. This indicates that, in the case of styryl molecules, the localization and excitation/emission properties can be optimized independently from each other and that finding mitochondrially targeted molecules that fluoresce at wavelengths >580 nm should be possible if a larger library were constructed.

To test how the measured spectral properties of the dyes in solution correspond to spectral properties of the dyes in living cells, UACC-62 melanoma cells were labeled with representative compounds in the library and visualized with an epifluorescence microscope. For excitation, filter sets were used to excite the dyes at three different wavelengths (405, 490, and 570 nm), and fluorescence was detected using a 500 nm dichroic and >510 long pass filter. Images were obtained from the cells at 200× magnification (Figure 8). As can be seen in the images, different dyes are optimally excited at different wavelengths, corresponding to the trends observed in the clustered, peak excitation plot. To illustrate this trend in the counterclockwise direction, the left bottom corner of the clustered emission graph corresponds to dyes that show the lowest excitation wavelength in the microscope images (405 nm). As one continues counterclockwise, the bottom right corner corresponds to dyes that excite at slightly higher wavelength (405/490 nm in the microscope images), while the upper right corresponds to dyes that excite at the highest wavelengths (490/570 nm in the microscope images). Continuing counterclockwise, as one moves toward the upper

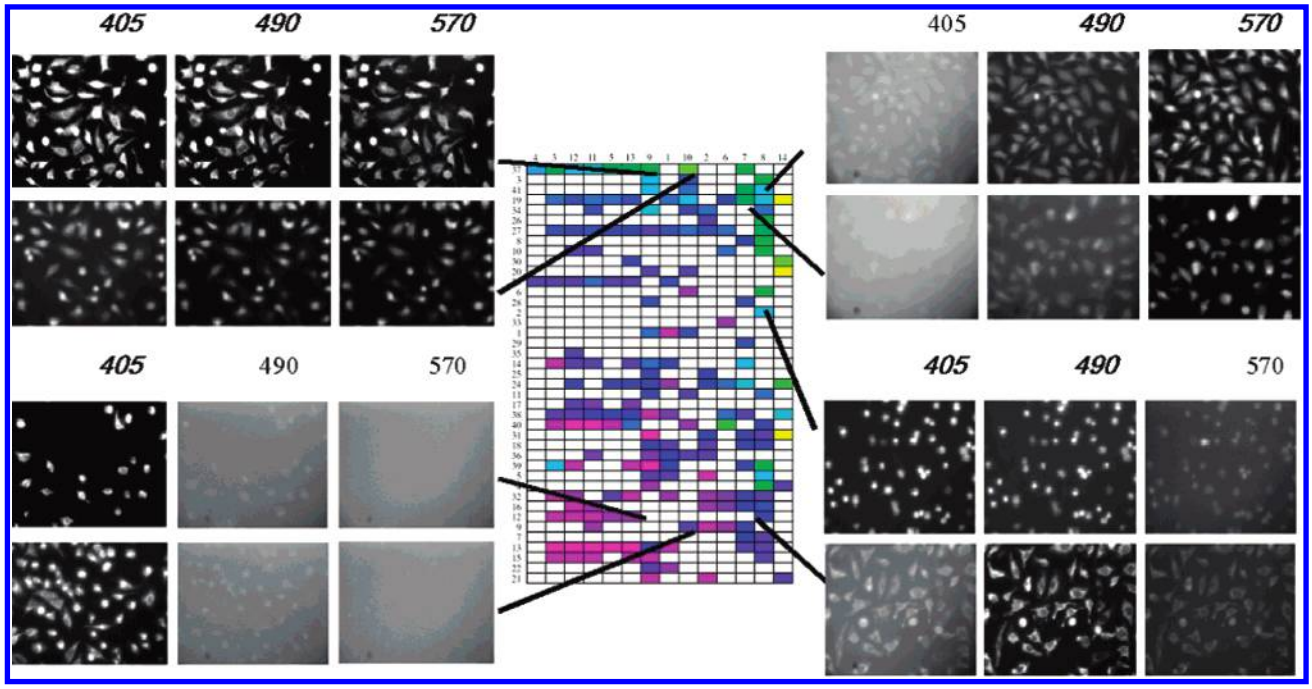

Figure 8. Epifluorescence microscopy analysis of selected styryl products selected from the excitation table (from Figure 4). Styryl products corresponding to A and B combinations yielding a range of peak emission wavelength were used to stain living cells and observed with various excitation filters (405, 490, and 570 nm), as indicated. Excitation wavelengths yielding the best fluorescence images are indicated in bolded letters.

left, dyes begin to excite at slightly lower wavelength in the microscope images (405/490/570 nm), and so on all the way back down to the lower left corner where dyes only excite at the lowest wavelengths.

General Survey of Structure−Property Relationships. Although a detailed quantum mechanical and experimental analysis of the relationship between the chemical structure of the molecules and their fluorescence properties is beyond the scope of this work, the overall trends in the data are consistent with expected relationships between styryl molecules' spectral properties and chemical structure. For the B groups, for example, B7, B8, B9, and B14 groups possess conjugated aromatic systems that contribute the greatest number of π electrons. In the spectral excitation and emission data (Figure 4), these groups strongly contribute toward the red end of the spectrum, which is expected based on quantum mechanical relationship between on the number of π electrons and the molecules higher excitation and emission. For the A groups, (N,N) dimethylaniline or an phenylamide substituent (A37, A27, A19, A18) contribute to increased resonance structures, as the partially pyramidal groups in aniline readily conjugate with the phenyl π system and lead to a delocalized positive charge spreading through the entire molecule[19] (Figure 9A). As expected in the excitation and emission data (Figure 4), these groups strongly contribute toward the red end of the spectrum, which is also expected based on quantum mechanical relationships based on the greater number of conjugated π electrons, and the greater degree of conjugation. In the case of nitrophenyl derivatives (A20, A21, A22, A36), the far red shift is not observed. This is expected because the oxygen groups are electron withdrawing[19] (Figure 9B), shifting fluorescence toward the blue.

In terms of the relationship between the chemical structure of styryl molecules and the localization properties, three alternative models can be proposed, for consideration (Figure 10). These models can be referred to as independent, cooperative, and noninteractive. According to the independent model, the A and B groups contribute to localization by virtue of their independent, isolatable interaction with different localization determinants localized in the organelle. According to the cooperative model, A and B groups contribute to localization by interaction with the same localization determinant in the organelle. These interactions may be partly isolatable, but the affinity is strongly dependent on both A and B being part of the same molecule. Last, according to the interactive model, A and B groups contribute to localization strictly as a result of how A and B interact when they are conjugated to each other, in a manner that A and B interaction with the organelle cannot be studied in isolation.

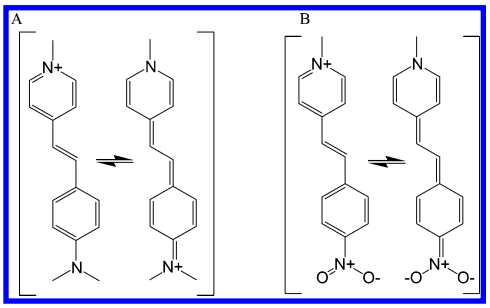

Figure 9. Resonance structure of (N, N) dimethylammonium phenyl (A) and nitrophenyl (B) styryl derivatives, illustrating charge delocalization and interactions between A and B moieties resulting from the conjugated, π electron system.

Among these three models, the independent model best accounts for the localization data obtained with the styryl library. Accordingly, the affinity of group B for the mitochondria can be added to the affinity of A with mitochondria, to determine the total affinity of the styryl molecule for mitochondria. As this may not be the only mechanism accounting for the independent relationship, we are currently

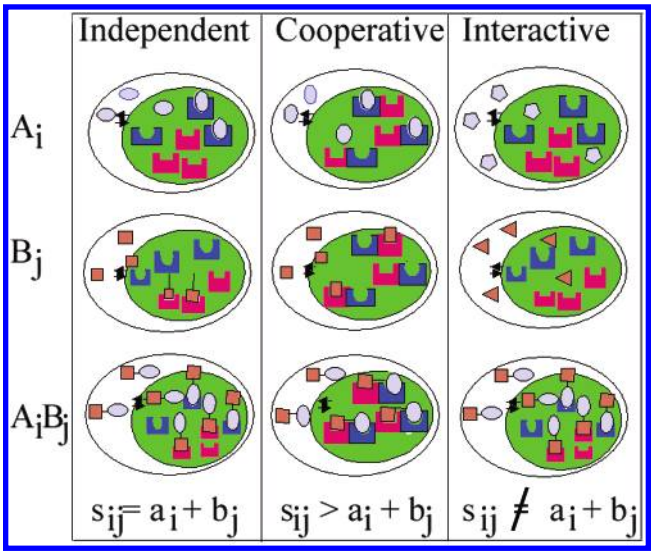

Figure 10. Illustration of three alternative hypothetical models that could explain mitochondrial localization. The A and B moieties are represented by geometrical shape (triangle, square, or otherwise). Mitochondria are represented by the inner green circle. Localization is ascribed to specific binding interaction between the A or B moieties and localization determinants present in the mitochondria.

performing experiments to test alternative mechanistic interpretations. Nevertheless, the additive decomposition analysis does suggest that cooperative and interactive binding interactions do not play a significant role in determining mitochondrial localization, across the entire library of styryl compounds. This is not intuitively obvious, because A and B moieties do indeed interact at the chemical level within the individual molecules: For example, the resonance structures exhibited by the molecules (Figure 9) suggest that the electron distribution across the entire molecule is strongly dependent on functional groups associated with A and B.

DISCUSSION

A combinatorial library of fluorescent compounds is constructed by coupling various combinations of moieties to a common fluorescent scaffold. One of the first examples of such combinatorial library is the synthesis of fluorescent, styryl compounds, combining an aldehyde drawn from a diverse set of 41 structures and a pyridinium drawn from a diverse set of 14 structures, condensed with each other with a secondary amine catalyst in 96 well plates.[17] Chemical properties (e.g. peak emission or excitation wavelength) and biological properties (e.g. subcellular localization) of the resulting styryl products can derive from characteristics already present in the individual building blocks that are used to synthesize the molecules or can emerge from complex physicochemical interactions observed only after the moieties are conjugated to each other. Nevertheless, because the individual building blocks are not fluorescent and the resulting product generally is, the resulting styryl compound can be readily detected and analyzed with a fluorometer and furthermore, purification is generally not required for preliminary analysis.[17]

Here, we report that the emerging fluorescence and localization properties of the styryl products are additively encoded in the structure of the constituent building blocks. Peak excitation and emission maxima, together with localization, are simply the sum of independent contributions of each of the two constituent moieties. Most of the functional moieties are associated with a specific and consistent influence on biological and chemical properties of compounds of which they are a part. This influence is largely independent of the structure of the remainder of the compound. In other words, a given A group may consistently be associated with redder emission peaks or with stronger mitochondrial localization, regardless of the B group to which it is joined and vice versa. This simple interaction allows for a comprehensive understanding of the variation of biological and chemical properties within the library, in terms of individual contributions of the constituent building blocks.

Practical Implications. Fluorescent biosensors can be useful experimental tools for cell biology, environmental monitoring, and, more recently, pharmaceutical screening applications.[20−29] Nevertheless, there are generally strict requirements in terms of what constitutes an ideal probe. For flow cytometry, for example, probes that are excited around 490 nm wavelength light are generally preferable, as flow cytometers commonly employ the 488 nm line of argon lasers as light source. For monitoring physiological function,[27,28,30−38] one may also want probes that are cell permeable, that associate with specific organelles, and that do not have a major effect on cell viability. For multiparameter cytometry,[39−48] probes that emit in narrow fluorescence bands at a variety of different wavelengths, that show reduced phototoxicity or bleaching, and that localize to a specific organelle may be highly desirable.

That simple, additive interactions in large part determine the spectral and localization characteristics in styryl dyes should facilitate the design and synthesis of styryl derivatives with ideal properties. For one, only a small fraction of the compounds need be synthesized and screened in order to attain accurate predictions of the localization and spectral properties throughout the library. Thus, for example, biased libraries that are optimally red and mitochondrial in localization or biased libraries that are optimally blue and nonmitochondrial in localization may be synthesized and screened, without having to synthesize and screen every possible styryl compound. Thus, a major practical application is to reduce the amount of screening that may be required to identify compounds in the library with optimal localization and spectral properties. Because the contribution of each of the two building blocks to localization and spectral properties are not interdependent, it should be possible to synthesize styryl derivatives optimized for localization and that span the visible spectrum in terms of excitation and emission peaks.

With respect to analysis and visualization of the data set, the present analysis allows reducing the dimensionality of the data and provides a natural, robust way of clustering and visualizing the data. In the past decade, the development of computational tools to handle massive amounts of data generated from high throughput screening experiments has been critical to the widespread adoption of combinatorial chemistry in drug discovery. Automated, quantitative analysis of structure−activity relationships (QSAR), together with data visualization tools, is essential for dealing with huge numbers of compounds. As a clustering and visualization method, the analysis here presented is ideally suited for classifying fluorescent, organelle-targeted molecular probes,

facilitating further synthesis, screening, and analysis of larger combinatorial libraries of fluorescent styryl molecules, for biosensor applications.

Mechanistic Inferences. While the statistical model used here is completely empirical, one can speculate as to the mechanistic nature of the relationships that we have found. For the spectral data, the additive relationship is plausible based on a simple "particle in a box" model, in which each of the two constituent moieties contributes a fixed number of π electrons to the styryl product. These π electrons resonate over the entire styryl structure via the conjugated bridge. Since the bridge is rigid, and the sizes of the moieties are roughly comparable, the "particle in a box" approximation explains the energy transitions in the product molecule as a noninteractive, additive function of characteristics (i.e. the number of π electrons and the physical dimensions of the space over which the electron resonates) contributed by each of the two moieties.

Additivity in subcellular localization could be explained as the sum of the chemical potential of the interactions, independently contributed by each of the two constituent moieties toward localization to a particular organelle. For interactions between the cationic B moieties and mitochondria, the electrostatic potential may be the primary determinant of mitochondrial localization, explaining the observed lack of differential influence of chemical diversity of the pyridinium/quinolinium group on localization. For the interaction between the lipophilic A moieties and mitochondria, this interaction may be a function of chemical potential of the A moiety across the mitochondrial membrane.

Interestingly, both of these inferences on the fluorescence and localization properties of the styryl compounds suggest experimental hypothesis that are testable under well-defined conditions. In the case of localization properties, the response of styryl molecules to a transmembrane potential can be accurately determined using liposomes in the presence of an ionic gradient and could be modeled using molecular dynamics simulations. In the case of spectral properties, quantum mechanical calculations may be used to independently establish how constituent building blocks of the styryl molecule contribute to the fluorescence properties of the resulting compounds. Experiments aimed at testing the various experimental hypotheses suggested by the analysis here presented are in progress.

ACKNOWLEDGMENT

We would like to thank Drs. J. Topliss and G. Crippen for insightful discussions and critical reading of the manuscript.

REFERENCES AND NOTES

(1) Chen, L. B. Mitochondrial membrane potential in living cells. *Annu. Rev. Cell Biol.* **1988**, *4*, 155−181.

(2) Lampidis, T. J.; Salet, C.; Moreno, G.; Chen, L. B. Effects of the mitochondrial probe rhodamine 123 and related analogues on the function and viability of pulsating myocardial cells in culture. *Agents Actions* **1984**, *14*, 751−757.

(3) Johnson, L. V.; Walsh, M. L.; Chen, L. B. Localization of mitochondria in living cells with rhodamine 123. *Proc. Natl. Acad. Sci. U.S.A.* **1980**, *77*, 990−994.

(4) Johnson, L. V.; Walsh, M. L.; Bockus, B. J.; Chen, L. B. Monitoring of relative mitochondrial membrane potential in living cells by fluorescence microscopy. *J. Cell Biol.* **1981**, *88*, 526−535.

(5) Smiley, S. T.; Reers, M.; Mottola-Hartshorn, C.; Lin, M.; Chen, A.; Smith, T. W.; Steele, G. D., Jr.; Chen, L. B. Intracellular heterogeneity in mitochondrial membrane potentials revealed by a J-aggregate-forming lipophilic cation JC-1. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 3671−3675.

(6) Mewes, H. W.; Rafael, J. The 2-(dimethylaminostyryl)-1-methylpyridinium cation as indicator of the mitochondrial membrane potential. *FEBS Lett.* **1981**, *131*, 7−10.

(7) Bereiter-Hahn, J. Dimethylaminostyrylmethylpyridiniumiodine (daspmi) as a fluorescent probe for mitochondria in situ. *Biochim. Biophys. Acta* **1976**, *423*, 1−14.

(8) Bereiter-Hahn, J.; Seipel, K. H.; Voth, M.; Ploem, J. S. Fluorimetry of mitochondria in cells vitally stained with DASPMI or rhodamine 6 GO. *Cell Biochem. Funct.* **1983**, *1*, 147−155.

(9) Snyder, D. S.; Small, P. L. Staining of cellular mitochondria with LDS-751. *J. Immunol. Methods* **2001**, *257*, 35−40.

(10) Bunting, J. R.; Phan, T. V.; Kamali, E.; Dowben, R. M. Fluorescent cationic probes of mitochondria. Metrics and mechanism of interaction. *Biophys. J.* **1989**, *56*, 979−993.

(11) Zoratti, M.; Favaron, M.; Pietrobon, D.; Petronilli, V. Nigericin-induced transient changes in rat-liver mitochondria. *Biochim. Biophys. Acta* **1984**, *767*, 231−239.

(12) Nicholls, D.; Ferguson, S. *Bioenergetics 2*; Academic Press: London, 1992.

(13) Modica-Napolitano, J. S.; Aprille, J. R. Delocalized lipophilic cations selectively target the mitochondria of carcinoma cells. *Adv. Drug Deliv. Rev.* **2001**, *49*, 63−70.

(14) Zhang, H.; Huang, H. M.; Carson, R. C.; Mahmood, J.; Thomas, H. M.; Gibson, G. E. Assessment of membrane potentials of mitochondrial populations in living cells. *Anal. Biochem.* **2001**, *298*, 170−180.

(15) Reed, P. W. Ionophores. *Methods Enzymol.* **1979**, *55*, 435−454.

(16) Rosania, G. R. Supertargeted Chemistry: Identifying Relationships Between Molecular Structures and their Sub-Cellular Distribution. *Curr. Top. Med. Chem.* **2003**, *3*, 659−685.

(17) Rosania, G. R.; Lee, J. W.; Ding, L.; Yoon, H. S.; Chang, Y. T. Combinatorial approach to organelle-targeted fluorescent library based on the styryl scaffold. *J. Am. Chem. Soc.* **2003**, *125*, 1130−1131.

(18) Lakowicz, J. R. *Principles of fluorescence spectroscopy*, 2nd ed.; Kluwer Academic/Plenum: New York, 1999; xxiii, 698.

(19) Streitwieser, A.; Heathcock, C. H. *Introduction to organic chemistry*, 3rd ed.; Macmillan: Collier Macmillan: New York, London, 1985; xviii, 1197.

(20) Zhang, J.; Campbell, R. E.; Ting, A. Y.; Tsien, R. Y. Creating new fluorescent probes for cell biology. *Nat. Rev. Mol. Cell Biol.* **2002**, *3*, 906−918.

(21) Ting, A. Y.; Kain, K. H.; Klemke, R. L.; Tsien, R. Y. Genetically encoded fluorescent reporters of protein tyrosine kinase activities in living cells. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 15003−15008.

(22) Rizzuto, R.; Brini, M.; De Giorgi, F.; Rossi, R.; Heim, R.; Tsien, R. Y.; Pozzan, T. Double labeling of subcellular structures with organelle-targeted GFP mutants in vivo. *Curr. Biol.* **1996**, *6*, 183−188.

(23) Miyawaki, A.; Llopis, J.; Heim, R.; McCaffery, J. M.; Adams, J. A.; Ikura, M.; Tsien, R. Y. Fluorescent indicators for Ca2+ based on green fluorescent proteins and calmodulin. *Nature* **1997**, *388*, 882−887.

(24) Gonzalez, J. E.; Tsien, R. Y. Voltage sensing by fluorescence resonance energy transfer in single cells. *Biophys. J.* **1995**, *69*, 1272−1280.

(25) Gonzalez, J. E.; Tsien, R. Y. Improved indicators of cell membrane potential that use fluorescence resonance energy transfer. *Chem. Biol.* **1997**, *4*, 269−277.

(26) Burdette, S. C.; Walkup, G. K.; Spingler, B.; Tsien, R. Y.; Lippard, S. J. Fluorescent sensors for Zn(2+) based on a fluorescein platform: synthesis, properties and intracellular distribution. *J. Am. Chem. Soc.* **2001**, *123*, 7831−7841.

(27) Taylor, D. L.; Woo, E. S.; Giuliano, K. A. Real-time molecular and cellular analysis: the new frontier of drug discovery. *Curr. Opin. Biotechnol.* **2001**, *12*, 75−81.

(28) Giuliano, K. A.; Taylor, D. L. Fluorescent-protein biosensors: new tools for drug discovery. *Trends Biotechnol.* **1998**, *16*, 135−140.

(29) Conway, B. R.; Minor, L. K.; Xu, J. Z.; Gunnet, J. W.; DeBiasio, R.; D'Andrea, M. R.; Rubin, R.; Giuliano, K.; DeBiasio, L.; Demarest, K. T. Quantification of G-Protein Coupled Receptor Internatilization Using G-Protein Coupled Receptor-Green Fluorescent Protein Conjugates with the ArrayScantrade mark High-Content Screening System. *J. Biomol. Screen* **1999**, *4*, 75−86.

(30) Waggoner, A.; DeBiasio, R.; Conrad, P.; Bright, G. R.; Ernst, L.; Ryan, K.; Nederlof, M.; Taylor, D. Multiple spectral parameter imaging. *Methods Cell. Biol.* **1989**, *30*, 449−478.

(31) Taylor, D. L.; Salmon, E. D. Basic fluorescence microscopy. *Methods Cell. Biol.* **1989**, *29*, 207−237.

(32) Post, P. L.; DeBiasio, R. L.; Taylor, D. L. A fluorescent protein biosensor of myosin II regulatory light chain phosphorylation reports a gradient of phosphorylated myosin II in migrating cells. *Mol. Biol. Cell* **1995**, *6*, 1755−1768.

(33) Giuliano, K. A.; Post, P. L.; Hahn, K. M.; Taylor, D. L. Fluorescent protein biosensors: measurement of molecular dynamics in living cells. *Annu. Rev. Biophys. Biomol. Struct.* **1995**, *24*, 405−434.
(34) Galbraith, W.; Wagner, M. C.; Chao, J.; Abaza, M.; Ernst, L. A.; Nederlof, M. A.; Hartsock, R. J.; Taylor, D. L.; Waggoner, A. S. Imaging cytometry by multiparameter fluorescence. *Cytometry* **1991**, *12*, 579−596.
(35) Farkas, D. L.; Baxter, G.; DeBiasio, R. L.; Gough, A.; Nederlof, M. A.; Pane, D.; Pane, J.; Patek, D. R.; Ryan, K. W.; Taylor, D. L. Multimode light microscopy and the dynamics of molecules, cells, and tissues. *Annu. Rev. Physiol.* **1993**, *55*, 785−817.
(36) DeBiasio, R.; Bright, G. R.; Ernst, L. A.; Waggoner, A. S.; Taylor, D. L. Five-parameter fluorescence imaging: wound healing of living Swiss 3T3 cells. *J. Cell Biol.* **1987**, *105*, 1613−1622.
(37) Bright, G. R.; Whitaker, J. E.; Haugland, R. P.; Taylor, D. L. Heterogeneity of the changes in cytoplasmic pH upon serum stimulation of quiescent fibroblasts. *J. Cell Physiol.* **1989**, *141*, 410−419.
(38) Bright, G. R.; Fisher, G. W.; Rogowska, J.; Taylor, D. L. Fluorescence ratio imaging microscopy. *Methods Cell Biol.* **1989**, *30*, 157−192.
(39) Smolewski, P.; Grabarek, J.; Kamentsky, L. A.; Darzynkiewicz, Z. Bivariate analysis of cellular DNA versus RNA content by laser scanning cytometry using the product of signal subtraction (differential fluorescence) as a separate parameter. *Cytometry* **2001**, *45*, 73−78.
(40) Smolewski, P.; Bedner, E.; Du, L.; Hsieh, T. C.; Wu, J. M.; Phelps, D. J.; Darzynkiewicz, Z. Detection of caspases activation by fluorochrome-labeled inhibitors: Multiparameter analysis by laser scanning cytometry. *Cytometry* **2001**, *44*, 73−82.
(41) Schmid, I.; Dagarag, M. D.; Hausner, M. A.; Matud, J. L.; Just, T.; Effros, R. B.; Jamieson, B. D. Simultaneous flow cytometric analysis of two cell surface markers, telomere length, and DNA content. *Cytometry* **2002**, *49*, 96−105.
(42) Poot, M.; Pierce, R. H. Detection of changes in mitochondrial function during apoptosis by simultaneous staining with multiple fluorescent dyes and correlated multiparameter flow cytometry. *Cytometry* **1999**, *35*, 311−317.
(43) Michalet, X.; Kapanidis, A. N.; Laurence, T.; Pinaud, F.; Doose, S.; Pflughoefft, M.; Weiss, S. The Power and Prospects of Fluorescence Microscopies and Spectroscopies. *Annu. Rev. Biophys. Biomol. Struct.* **2003**.
(44) Lee, G. S.; Ryu, K. S.; Rha, J. G.; Kim, S. P.; Namkoong, S. E.; Han, K. T. Multiparametric flow cytometric analysis in a breast cancer cell line (MCF-7). *J. Obstet. Gynaecol. Res.* **2002**, *28*, 141−148.
(45) Gordon, K. M.; Duckett, L.; Daul, B.; Petrie, H. T. A simple method for detecting up to five immunofluorescent parameters together with DNA staining for cell cycle or viability on a benchtop flow cytometer. *J. Immunol. Methods* **2003**, *275*, 113−121.
(46) Corver, W. E.; Koopman, L. A.; van der Aa, J.; Regensburg, M.; Fleuren, G. J.; Cornelisse, C. J. Four-color multiparameter DNA flow cytometric method to study phenotypic intratumor heterogeneity in cervical cancer. *Cytometry* **2000**, *39*, 96−107.
(47) Perez, O. D.; Nolan, G. P. Simultaneous measurement of multiple active kinase states using polychromatic flow cytometry. *Nat. Biotechnol.* **2002**, *20*, 155−162.
(48) Herzenberg, L. A.; Parks, D.; Sahaf, B.; Perez, O.; Roederer, M. The history and future of the fluorescence activated cell sorter and flow cytometry: a view from Stanford. *Clin. Chem.* **2002**, *48*, 1819−1827.

CI0341215