Title

Fluorescence-aided molecule sorting: Analysis of structure and interactions by alternating-laser excitation of single molecules

Journal

Proceedings of the National Academy of Sciences of the United States of America, 101(24)

ISSN

0027-8424

Authors

Kapanidis, A N
Lee, N K
Laurence, T A
et al.

Publication Date

2004-06-01

Peer reviewed

Fluorescence-aided molecule sorting: Analysis of structure and interactions by alternating-laser excitation of single molecules

Achillefs N. Kapanidis*†, Nam Ki Lee*‡, Ted A. Laurence, Sören Doose, Emmanuel Margeat, and Shimon Weiss†

Departments of Chemistry and Biochemistry and of Physiology, University of California, 607 Charles E. Young Drive East, Los Angeles, CA 90095

Edited by Robin M. Hochstrasser, University of Pennsylvania, Philadelphia, PA, and approved April 22, 2004 (received for review March 11, 2004)

We use alternating-laser excitation to achieve fluorescence-aided molecule sorting (FAMS) and enable simultaneous analysis of biomolecular structure and interactions at the level of single molecules. This was performed by labeling biomolecules with fluorophores that serve as donor–acceptor pairs for Förster resonance energy transfer, and by using alternating-laser excitation to excite directly both donors and acceptors present in single diffusing molecules. Emissions were reduced to the distance-dependent ratio *E*, and a distance-independent, stoichiometry-based ratio *S*. Histograms of *E* and *S* sorted species based on the conformation and association status of each species. *S* was sensitive to the stoichiometry and relative brightness of fluorophores in single molecules, observables that can monitor oligomerization and local-environment changes, respectively. FAMS permits equilibrium and kinetic analysis of macromolecule-ligand interactions; this was validated by measuring equilibrium and kinetic dissociation constants for the interaction of *Escherichia coli* catabolite activator protein with DNA. FAMS is a general platform for ratiometric measurements that report on structure, dynamics, stoichiometries, environment, and interactions of diffusing or immobilized molecules, thus enabling detailed mechanistic studies and ultrasensitive diagnostics.

single-molecule fluorescence spectroscopy | Förster resonance energy transfer | biomolecular interactions | catabolite activator protein | protein–DNA interactions

Understanding biological mechanisms requires analysis of biomolecular structure and interactions, as well as monitoring of their changes as a function of time. Presently, few methods are robust enough to analyze structure and interactions simultaneously. One of these methods is Förster resonance energy transfer (FRET), a method based on the nonradiative transfer of excitation energy from a donor (D) fluorophore to a complementary acceptor (A) fluorophore (1). The FRET efficiency E is a sensitive function of D–A distance R, because $E = [1 + (R/R_o)^6]^{-1}$ (where R_o is a constant that equals the D–A distance at $E = 50\%$), allowing use of FRET as a "spectroscopic ruler" for the 1- to 10-nm scale. Moreover, presence of intermolecular FRET allows detection and analysis of molecular interactions.

Often, FRET analysis is complicated by heterogeneity (because of free, unlabeled, inactive, or unsynchronized species) inherent in ensembles of biomolecules containing several polypeptide chains and/or nucleic acids. This is addressed in part by measuring FRET at the level of single molecules (single-pair FRET or spFRET; refs. 2 and 3). Using single-laser excitation for spFRET on diffusing molecules, the donor is excited directly during the transit of a molecule through an observation volume defined by a focused laser beam and confocal optics (green oval, Fig. 1*A*). When an acceptor is close to the donor, part of the donor-excitation energy is transferred to the acceptor, which emits at wavelengths longer than the donor. When plotted as a function of time, the emissions for each fluorophore appear as "bursts" of fluorescence (green and red curves in Fig. 1*A*); ratios of the emissions report on D–A distance (3) (Fig. 1*A*, a1). spFRET has been used to study dynamics of proteins, nucleic acids, and their complexes (4–6).

However, spFRET is not a general platform for quantitative analysis of structure; it has mainly been used to identify distance changes and their kinetics. This is due to the numerous corrections required to measure accurate FRET efficiencies within single molecules, and the presence of chemically or photophysically induced species that obscure FRET measurements when $R > 6–8$ nm.

Moreover, spFRET is not a general platform for quantitative analysis of molecular interactions. For example, for the interaction $M^A + L^D \rightleftharpoons M^AL^D$ (where M^A is an A-labeled macromolecule, and L^D is a D-labeled ligand), spFRET cannot quantitate accurately the major species involved in the interaction. First, spFRET yields a measurable signal only when D–A distances in the M^AL^D complex are sufficiently short (typically $R_{D–A} < 6–8$ nm) to distinguish complexes from free L^D species (Fig. 1*A*; a1 vs. a3). This proximity constraint is difficult to satisfy in all cases, especially for large complexes or interacting proteins of unknown structure. Second, inactive states of FRET acceptors result in ML^D species that exhibit D-only characteristics (3), leading to apparent increases in the free L^D species. Third, no M^A species are detected (Fig. 1*A*, a4), because direct A-excitation at the wavelength of D-excitation is minimized to reduce crosstalk. Fourth, complexes with stoichiometries other than 1:1 (e.g., $M^A[L^D]_2$) cannot be identified by spFRET.

The cumulative effect of such limitations, combined with complications caused by substoichiometric labeling, fluorophore photophysics, photobleaching, and aggregation or dissociation phenomena have prevented the full realization of the spFRET potential.

Here, we introduce a single-molecule fluorescence spectroscopy that serves as a general tool for monitoring structure and interactions simultaneously. This is achieved by obtaining D-excitation and A-excitation-based observables (Fig. 1) for each single molecule by using an alternating-laser excitation (ALEX) scheme, during which we switch rapidly between a D-excitation and an A-excitation laser (Fig. 1*C*). This scheme recovers distinct emission signatures for all diffusing species (Fig. 1) by calculating two fluorescence ratios: the FRET efficiency E (3, 7, 8), which reports on D–A distance in the M^AL^D complex, and the distance-independent ratio S, which reports on the D–A stoichiometry of all species. S provides important information even in the absence of close proximity between fluorophores; it allows thermodynamic and kinetic analysis of inter-

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: ALEX, alternating-laser excitation; A, acceptor; D, donor; FRET, Förster resonance energy transfer; FAMS, fluorescence-aided molecule sorting; CAP, catabolite activator protein; spFRET, single-pair FRET; TMR, tetramethylrhodamine.

*A.N.K. and N.K.L. contributed equally to this work.

†To whom correspondence may be addressed. E-mail: kapanidi@chem.ucla.edu or sweiss@chem.ucla.edu.

‡Present address: School of Chemistry, Seoul National University, Seoul, Korea 151-747.

www.pnas.org/cgi/doi/10.1073/pnas.0401690101

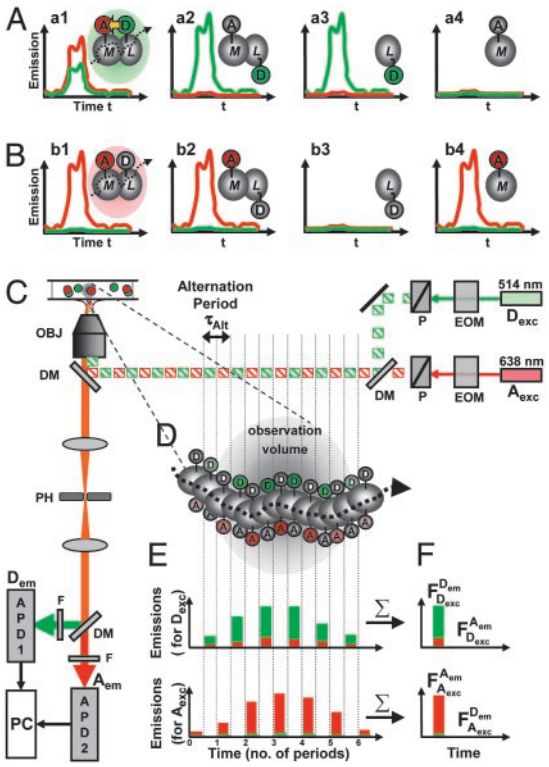

Fig. 1. ALEX allows detection of D-excitation- and A-excitation-based emissions for single diffusing molecules, enabling sorting. *M*, macromolecule; *L*, ligand. (*A*) Emissions caused by D-excitation (green oval). Short D–A distances in M^AL^D (a1) result in high FRET, detected as coincident bursts of fluorescence with high A-emission (red curve) and low D-emission (green curve). However, there is no discrimination between low-FRET M^AL^D (a2) and free L^D (a3); free M^A are undetected (a4). (*B*) Emissions caused by A-excitation (red oval). Direct A-excitation results in high A-emission when A is present (b1, b2, and b4). (*C*) ALEX microscopy. EOM, electrooptical modulator; P, polarizer; DM, dichroic mirror; OBJ, objective; PH, pinhole; F, filter; APD, avalanche photodiode. When EOM-polarizer combinations are used, lasers yield an excitation alternating between 514-nm (D_{exc}) and 638-nm (A_{exc}) light (green and red hatched boxes) with period τ_{alt}. The alternating laser excites the sample, which emits fluorescence (orange line) collected in D-emission (D_{em}) and A-emission (A_{em}) channels. (*D*) Probing D-excitation- and A-excitation-based emissions for a diffusing molecule. A low-*E* D–A species crosses the observation volume and, upon excitation, generates a fluorescence burst. (*E*) When the 514-nm excitation is on, D-excitation-based D- and A-emissions (green and red columns, respectively) are collected; when the 638-nm excitation is on, A-excitation-based emissions are collected. (*F*) Summing the four excitation/emission streams yields four emissions for a single burst, allowing calculation of *E* and *S*, and enabling sorting.

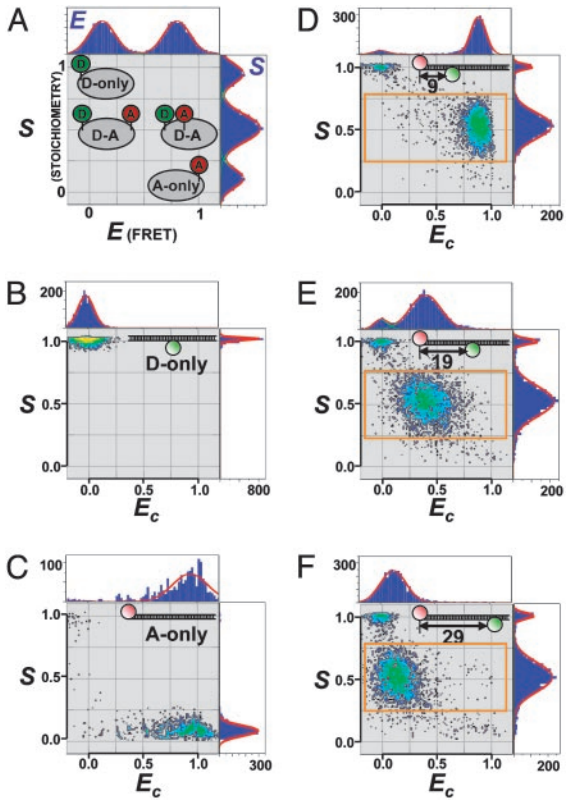

Fig. 2. Sorting single molecules using FAMS. (*A*) *E-S* histogram for D-only, A-only, and D–A species with different R_{D-A}. *E* (or E_c) sorts species according to FRET and R_{D-A}, reporting on structure; *S* sorts species according to D–A stoichiometry, reporting on interactions. Sorting is also possible by using 1D E_c or *S* histograms (in blue; red line, sum of Gaussian fits; green line, individual Gaussian fits). (*B*) D-only DNA. For all fragments, [DNA] = 50 pM; τ_{alt} = 50 μs. (*C*) A-only DNA. (*D*) High-*E* DNA. (*E*) Intermediate-*E* DNA. (*F*) Low-*E* DNA.

actions, identification of interaction stoichiometry, and study of local environment (as detected by changes in the fluorophore brightness). Combination of *E* and *S* on 2D histograms (Fig. 2*A*) allows virtual sorting of single molecules (9); we define this analysis as fluorescence-aided molecule sorting (FAMS), and we designate its implementation using alternating-laser excitation as ALEX–FAMS. ALEX–FAMS is a homogeneous, "mix-and-read" assay in which interacting species are combined and optical readouts report simultaneously on their association status and conformational status. We demonstrate the capabilities of ALEX–FAMS by studying fluorescent DNA fragments, and the interaction of *Escherichia coli* catabolite activator protein (CAP) with DNA. ALEX offers a general platform for generating excitation-based observables that report on biomolecular structure and interactions, and because it can monitor the kinetics of assembly/disassembly processes and conformational transitions, it may allow real-time observations of complete mechanisms at the level of single molecules.

Materials and Methods

Principle of ALEX. ALEX is based on alternation between lasers (Fig. 1 *C* and *D*) at time scales >2-fold faster than the ≈1-ms diffusion time of a fluorescent molecule through a femtoliter detection volume.

The first laser (D_{exc}) excites the donor directly, and can excite the acceptor indirectly if R_{D-A} is within FRET range (Fig. 1*A*). After dividing the time axis into intervals *i* of durations matching the alternation period τ_{alt}, we define the background-corrected D-excitation-based D-emission as $f_{D_{exc}}^{D_{em}}$, and the D-excitation-based A-emission as $f_{D_{exc}}^{A_{em}}$ (Fig. 1*D*). A fluorescence burst (corresponding to a single diffusing molecule) is defined as the sum of photons emitted in a contiguous series of alternation periods, where $i = i_0$ to $i_0 + n - 1$, and *n* is the number of periods in a single burst. The D-excitation-based emissions during a single burst are

$$F_{D_{exc}}^{D_{em}} = \sum_{i=i_0}^{i_0+n-1} f_{D_{exc}}^{D_{em}}(i) \qquad F_{D_{exc}}^{A_{em}} = \sum_{i=i_0}^{i_0+n-1} f_{D_{exc}}^{A_{em}}(i). \qquad [1]$$

These emissions report on the D–A distance for the burst-generating molecule through the calculation of ratio *E* (FRET efficiency) for a single burst

$$E = F_{D_{exc}}^{A_{em}} / (F_{D_{exc}}^{A_{em}} + \gamma F_{D_{exc}}^{D_{em}}), \qquad [2]$$

where $\gamma = (\phi_A \eta_A)/(\phi_D \eta_D)$ is a detection correction factor that depends on donor and acceptor quantum yields ϕ_D and ϕ_A, and detection efficiencies η_D and η_A of donor and acceptor emission

channels; typically, $0.5 < \gamma < 2$. [Exact knowledge of γ is not important for FAMS, but is important for accurate E and R_{D-A} measurements (N.K.L., A.N.K., and S.W., unpublished data); for most of this work, $\gamma \approx 1$, allowing accurate E measurements.] Ratio E ($0 \le E \le 1$) assumes very low values for D-only species, low values for D–A species with $R_{D-A} > R_o$, and high values for D–A species with $R_{D-A} < R_o$ (Fig. 1*A*, a1–a3). Because the acceptor is not excited by D-excitation, discrete values of E for A-only species are observed (Fig. 1*A*, a4). Because E depends only on D-excitation-based emissions, it is independent of the alternation period or duty cycle. Moreover, E contains R_{D-A} information only for species with active donors and acceptors.

The second laser (A_{exc}) excites the acceptor directly and does not excite the donor, allowing the formulation of the ALEX-based ratio S that reports on D–A stoichiometry. Similarly to D-excitation based observables, we define the background-corrected A-excitation-based D-emission as $f^{D_{em}}_{A_{exc}}$, and the A-excitation-based A-emission as $f^{A_{em}}_{A_{exc}}$. For single bursts, we define the A-excitation-based A-emission $F^{A_{em}}_{A_{exc}}$ and the A-excitation-based D-emission $F^{D_{em}}_{A_{exc}}$ as:

$$F^{A_{em}}_{A_{exc}} = \sum_{i=i_0}^{i_0+n-1} f^{A_{em}}_{A_{exc}}(i) \qquad F^{D_{em}}_{A_{exc}} = \sum_{i=i_0}^{i_0+n-1} f^{D_{em}}_{A_{exc}}(i). \qquad [3]$$

For a single burst, the modified sum of D-excitation-based emissions is defined as $F_{D_{exc}} = F^{A_{em}}_{D_{exc}} + \gamma F^{D_{em}}_{D_{exc}}$, whereas the sum of A-excitation-based emissions is $F_{A_{exc}} = F^{A_{em}}_{A_{exc}} + F^{D_{em}}_{A_{exc}}$. We define S as

$$S = F_{D_{exc}}/(F_{D_{exc}} + F_{A_{exc}}). \qquad [4]$$

Use of S allows stoichiometry observations that are independent of the diffusion path (as in the case of E). Modified ratios that are related to S but emphasize specific aspects of stoichiometry can also be formulated; this was done previously (10, 11) but in the absence of FRET. S ($0 \le S \le 1$) assumes distinct values for all species in mixtures of interacting components (Fig. 2*A*). After adjusting the excitation to obtain $F_{D_{exc}} \approx F_{A_{exc}}$ for a D–A species (*Supporting Text*, which is published as supporting information on the PNAS web site), S for D-only species is high, ≈1 (because $F_{A_{exc}} = 0$), and S for A-only species is low, in the 0–0.2 range (because of low $F_{D_{exc}}$); S for D–A species characterized by any R_{D-A} assumes intermediate values, in the 0.3–0.8 range. The distance-independent nature of S is caused by the distance-independent nature of $F_{D_{exc}}$ and $F_{A_{exc}}$ (through use of detection–correction factor γ; *Supporting Text*), making S and E independent observables. S is sensitive to changes in the brightness of the fluorophores in D–A species, an ability that can probe changes in the local environment. Combination of E and S in 2D histograms enables FAMS and quantitation of sorted species (Fig. 2*A*), while maintaining D–A distance information.

DNA, CAP, and CAP–DNA Complexes. DNA fragments were prepared by using automated synthesis (12), labeled, and hybridized to form D-only, A-only, and D–A samples (*Supporting Text* and Fig. 7, which is published as supporting information on the PNAS web site). [C17;S178]CAP was a gift from Richard Ebright (12). CAP was site-specifically labeled with tetramethylrhodamine (TMR) on Cys-17 (*Supporting Text*). Labeling efficiency was 45%, and DNA-binding activity of CAP^{TMR} was 25%. To form CAP–DNA complexes, DNA and CAP were incubated for 15–60 min at 25°C in CAP-binding buffer (CBB; 20 mM Hepes-NaOH, pH 7/200 mM NaCl/1 mM DTT/1 mM mercaptoethylamine/100 μg/ml BSA/5% glycerol), with or without 0.2 mM cAMP.

Sample Preparation. DNA samples were diluted in SMF buffer (10 mM Hepes-NaOH, pH 7/500 mM NaCl/100 μg/ml BSA/1 mM mercaptoethylamine/5% glycerol); CAP–DNA complexes were diluted in CBB with or without cAMP. Final DNA concentrations were 10–50 pM, resulting in a low probability (<0.5%) of simultaneous presence of two molecules in the detection volume (≈3 fl) (3).

Data Acquisition and Analysis. A single-molecule fluorescence setup (3, 13) was modified to allow ALEX of diffusing species (Fig. 1*C*) using 514-nm light from an Ar^+ laser, and 638-nm light from a diode laser; the alternation period was 25–3,000 μs. Alternation was achieved by using electrooptical modulators (EOM) combined with polarizers. By rotating the polarization of each laser beam individually before directing it to the polarizer, the lasers were switched on or off. Extinction ratios (ratios of laser intensities when a laser is on or off) were >100:1 for each laser. The duty cycle for each laser was 38–49%; temporal crosstalk between excitations is eliminated by allowing a 3-μs interval between excitations (e.g., 47 μs 514-nm excitation, 3 μs no excitation, 47 μs 638-nm excitation, 3 μs no excitation). Before entering the objective, the beams were spatially filtered through a single-mode fiber and circularly polarized by using achromatic waveplates. The continuous-wave excitation intensities were $(3.3–16.7) \times 10^4$ W/cm^2 for 514-nm excitation (D_{exc}), and $(0.7–5.3) \times 10^4$ W/cm^2 for 638-nm excitation (A_{exc}).

For data analysis, photons detected at the donor or acceptor emission channel (Fig. 1 *D* and *E*) were assigned to either donor or acceptor excitation based on arrival time, generating emissions $f^{A_{em}}_{D_{exc}}$, $f^{D_{em}}_{D_{exc}}$, $f^{A_{em}}_{A_{exc}}$, and $f^{D_{em}}_{A_{exc}}$ (Fig. 1*E* and Fig. 8, which is published as supporting information on the PNAS web site). Signals caused by A-emission into the D-emission channel, and background due to scattering were negligible. Emissions were analyzed to identify fluorescence bursts (3) (*Supporting Text*), recovering $F^{A_{em}}_{D_{exc}}$, $F^{D_{em}}_{D_{exc}}$, $F^{A_{em}}_{A_{exc}}$, and $F^{D_{em}}_{A_{exc}}$, and allowing calculation of E and S for each molecule and construction of 2D E-S histograms (Fig. 2*A*). In specific cases, E was corrected for D-emission at the wavelength of A-emission by using the E of D-only species (3); the corrected value is referred to as E_c.

E distributions were unaffected by laser-alternation characteristics such as the excitation-intensity ratio or alternation period τ_{Alt} (Figs. 9*A* and 10, which are published as supporting information on the PNAS web site). On the other hand, S distributions depended on the excitation-intensity ratios (Fig. 9*B*). Moreover, the width of S distributions increased when τ_{Alt} exceeded the diffusion time τ_D for the species of interest (Fig. 10*B*). Hence, the working range of τ_{Alt} is defined by a lower τ_{Alt} limit set by the EOM response time (≈2 μs), and a higher τ_{Alt} limit set by diffusion time τ_D (for our setup, ≈400 μs for a 35-bp DNA fragment).

Results and Discussion

Sorting Species with Different D–A Stoichiometry and D–A Distance Using FAMS. To test the ability of FAMS to monitor D–A stoichiometry (and thus association status), we prepared DNA fragments that served as D-only, A-only, and D–A species (Fig. 2). To test the ability of FAMS to monitor D–A distance (and thus conformational status), we prepared three DNA fragments, each containing both a donor (TMR) and an acceptor (Alexa 647, A647), separated by 9, 19, or 29 bp and located on opposite phases of the DNA helix. This scheme results in attachment-point distances R_{att} of ≈4, ≈7, or ≈11 nm, corresponding to high-E, intermediate-E, and low-E DNA (Figs. 2 and 7; $R_{o,\ TMR\rightarrow A647} \approx 6.4$ nm for samples examined (12). To evaluate the accuracy of the extracted D–A distances, we used a corrected expression of E (E_c; *Materials and Methods*).

The E_c-S histogram for D-only and A-only samples revealed single species with the expected E_c and S; low E_c and high S for D-only DNA (Fig. 2*B* and Table 1), and high E_c and low S for A-only DNA (Fig. 2*C* and Table 1). In contrast, the E_c-S histogram for high-E DNA (Fig. 2*D*) revealed a major species with high E_c (because $R_{D-A} < R_o$) and intermediate S (because $F^{A_{em}}_{A_{exc}} \approx F^{D_{em}}_{D_{exc}}$), along with few D-only species (because of excess bottom strand used for hybridization, and to inactive acceptor; ref. 3), and very few A-only species. The 1D E_c histogram fits to a Gaussian distribution, with $E_c = 0.90 \pm 0.07$ (which reflects the mean ± standard

Table 1. *E* and *S* results for the main species of Fig. 2, along with comparison of measured interprobe distances $R_{D\text{-}A}$ with distances between points of attachment R_{att}

	E_c, mean ± 1σ	S, mean ± 1σ	R_{D-A}, nm	R_{att}, nm
D-only	−0.02 ± 0.07	1.01 ± 0.02		
A-only	0.94 ± 0.16	0.08 ± 0.04		
D-A, high FRET	0.90 ± 0.07	0.53 ± 0.10	4.4	3.6
D-A, intermediate FRET	0.40 ± 0.14	0.53 ± 0.12	6.9	7.0
D-A, low FRET	0.12 ± 0.12	0.51 ± 0.10	9.0	10.5

deviation σ, and not the error of the mean; the error of the mean is $\Delta E \approx 0.01$) corresponding to $R_{D-A} \approx 4.4$ nm, and consistent with $R_{att} \approx 4$ nm. The 1D S histogram also fits to a Gaussian distribution ($S = 0.53 \pm 0.10$). Similarly to E, part of the S width is statistical ("shot-noise"), because S is a ratio of fluorescence intensities with low photon counts. Moreover, because S is a ratio of three emissions (vs. two emissions for E), the relative contribution of statistical noise to the overall width is expected to be larger than for E. As in the case of E, other undefined factors contribute to the width of the distribution (3, 5).

Increasing D–A separation to 19 bp (Fig. 2*E*) decreases E_c for the D–A species, but leaves S unchanged (Table 1), showing that S and E_c are independent; E_c was consistent with the expected R_{att} of ≈7 nm. Increasing D–A separation to 29 bp (Fig. 2*F*) further decreases E_c for the D–A species, but does not affect S (Table 1). The recovered E_c value corresponds to a distance slightly shorter than expected for R_{att} ≈11 nm; the difference is mainly caused by absence of a correction for A-emission due to direct A-excitation (ΔE_c of 0.05–0.10), which becomes significant at $R_{D-A} \gg R_o$. This correction can be made by using the A-only species as an internal control, thus defining a general way for subtracting non-FRET contributions from the $F^{A_{em}}_{D_{exc}}$ signal (refs. 1 and 14 and N.K.L., A.N.K., and S.W., unpublished data). Overall, increasing R_{D-A} decreases E_c but leaves S unchanged; moreover, E values yield distance constraints consistent with B-DNA structure (3).

Sorting and selecting species (e.g., D–A species; in rectangles of Fig. 2 *D–F*) allows removal of D-only species that contaminate spFRET studies to variable degrees (3). In cases of fast dissociation of complexes of interest or low activity of interacting components, D-only species comprise > 90% of the detected species, thwarting any spFRET analysis; this problem is more acute for D–A species with $R_{D-A} \gg R_o$. Removing D-only species allows observation of D–A species with low E, increasing both the dynamic range and accuracy of the distances constraints. Moreover, because even D–A species with $E = 0$ are identified as "bound," no proximity is required to identify interactions by FAMS.

Quantitative Analysis of Mixtures. To show that FAMS can analyze quantitatively mixtures of species with comparable diffusion times, we examined mixtures of low-E and high-E DNA (Fig. 3 *A–C*). FAMS of the 1:1 mixture sorted the D–A species of the DNA components along the E axis (Fig. 3*A*), yielding concentrations and E distributions for the mixture components that were within 5% of the concentrations and distributions obtained for the pure components. Similar results were obtained for 1:2 and 2:1 low-E/high-E DNA mixtures (Fig. 3 *B* and *C*), validating the quantitative nature of FAMS, which is further supported by the linear relation of number-of-bursts and analyte concentration for the 0–300 pM range (Fig. 11, which is published as supporting information on the PNAS web site).

FAMS Can Detect Dimerization. Oligomerization is a common modulator of protein function. Because S is sensitive to the ratio of donors/acceptors per molecule [e.g., for species D_2-A, D–A, and D–A_2, $S_{(D2\text{-}A)} > S_{(D\text{-}A)} > S_{(D\text{-}A2)}$], it can be used for monitoring

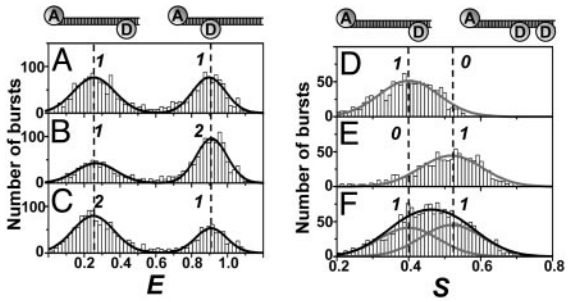

Fig. 3. FAMS can analyze mixtures of samples with different E or D–A stoichiometry. Black lines, sum of Gaussian fits; gray lines, individual Gaussian fits; dashed lines, means of individual Gaussian fits. (*A–C*) E histograms for mixtures of samples with different E. Histograms are for D–A species sorted from E-S histograms (using $0.3 < S < 0.8$). Analysis of 1:1 (*A*), 1:2 (*B*), and 2:1 (*C*) mixtures of low-E/high-E DNA yields ratios of 100:97, 100:206, and 100:54 for low-E/high-E species, respectively. (*D–F*) S histograms for D–A species with different D–A stoichiometries. (*D*) D–A species. (*E*) D_2-A species. (*F*) 1:1 mixture of D–A:D_2-A species.

oligomerization. We thus compared a DNA carrying one donor (Alexa 532, A532) and one acceptor (Alexa 647), with a DNA carrying two donors and one acceptor. The E-S histograms for the two DNA fragments were distinct (Fig. 12 *A* and *B*, which is published as supporting information on the PNAS web site), mainly along S (Fig. 3 *D* and *E*); as expected, $S_{(D2\text{-}A)} > S_{(D\text{-}A)}$ (0.53 ± 0.08 vs. 0.40 ± 0.07). Analysis of a 1:1 mixture of D_2-A and D–A yielded an E-S histogram (Fig. 12*C*) having a single D–A peak with a wider S distribution (0.46 ± 0.11; Fig. 3*F*), consistent with two species with closely spaced S distributions. Using a double fixed-Gaussian fit (with individual distributions described by using means and standard deviations of pure D_2-A and D–A), we recovered a D–A/D_2–A ratio of 0.9:1. Modified stoichiometry ratios, such as $S_{mod} = F_{D_{exc}}/F_{A_{exc}}$ (similar to ref. 11) increases the resolution between the D–A species of interest. Hence, FAMS of the monomeric and dimeric forms of a dimerizing system can extract the ratio of D–A and D_2-A species as a function of the concentration of D-labeled molecule, and yield dimerization constants; this can be performed either in the absence or presence of FRET, something not possible with spFRET.

FAMS Can Sense Fluorophore Brightness. The brightness of a fluorophore is the average emission rate per molecule, and it is a function of the efficiency of excitation and emission processes (13). S is sensitive to the relative brightness of the fluorophores in D–A species; such a property might be useful for monitoring changes in the local environment of the fluorophores. For example, if a conformational change alters the local environment of a FRET donor leading to quenching (translating to lower $F_{D_{exc}}$ in Eq. **4**), it leads to a decrease in S. This signal is independent of FRET; in fact, it allows corrections that differentiate E changes caused by distance changes from E changes caused by changes in the characteristic distance R_o for the given D–A pair, something that cannot be easily performed by spFRET.

To demonstrate that FAMS can sense relative brightness, we prepared two DNA fragments, each with a different donor (TMR or Alexa532) but with an identical acceptor (Alexa 647), and identical D–A distance (Fig. 4). Because of its spectral properties, Alexa532 is brighter than TMR for the selected combination of excitation and emission wavelengths (mainly because of larger extinction coefficient at the excitation wavelength); however, TMR transfers energy more efficiently to the acceptor (because $R_{o,A532\rightarrow A647} < R_{o,TMR\rightarrow A647}$). Indeed, use of TMR results in D–A species with the highest E and lowest S (Fig. 4*A*; $E \approx 0.45$, $S \approx 0.35$). Switching donor to Alexa 532 (Fig. 4*B*) decreases E, increases S, and decreases both E and S widths

BIOPHYSICS

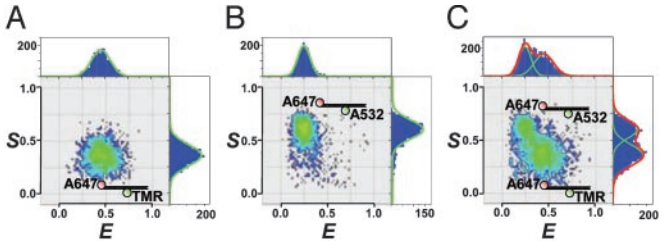

Fig. 4. FAMS is sensitive to fluorophore brightness. *E-S* histograms for D–A species with different donor but identical acceptor; D–A species were selected as described in *Supporting Text*. (*A*) $DNA^{TMR/A647}$. (*B*) $DNA^{A532/A647}$; both *E* and *S* change. (*C*) A 1:1 mixture of $DNA^{TMR/A647}$ and $DNA^{A532/A647}$; fitting along *E* or *S* recovers concentration ratios within 10% of the predicted 1:1 ratio.

(because of lower statistical noise). Fitting the *E* or *S* histogram for a 1:1 mixture of $DNA^{TMR/A647}$ with $DNA^{A532/A647}$ (Fig. 4*C*) recovers concentrations within 10% of the concentrations obtained for individual DNAs. The ability to sense fluorophore brightness turns FAMS into the single-molecule counterpart of popular ensemble–fluorescence assays that monitor protein conformational changes through changes in the brightness of environmentally sensitive probes (such as tryptophan and dansyl).

Macromolecule–Ligand Interactions. Because FAMS identifies D-only, A-only, and D–A species, it is suitable for analysis of interactions. D-only and A-only species represent free interactants, and D–A represents a complex (Fig. 5*A*); equilibrium binding and kinetic rate constants can be measured by simply counting molecules of the three major species. For a macromolecule-ligand interaction with equilibrium dissociation constant K_d, the fraction of bound macromolecules (fractional occupancy θ) at ligand concentration $[L]$ is $\theta = [ML]/([ML] + [M]) = [L]/([L] + K_d)$ (Langmuir–Hill equation, ref. 15). For macromolecule M^A and ligand L^D (Fig. 5*A*):

$$\theta = [D - A]/([D - A] + [A\text{-}only])$$
$$= [D\text{-}only]/([D\text{-}only] + K_d) \quad [5]$$

θ is extracted reliably from *E-S* histograms (*Supporting Text*). Measuring θ at several ligand concentrations $[L^D]$ allows calculation of K_d; θ can also be monitored as a function of time to evaluate association/dissociation kinetics. FAMS-based interaction analysis is unaffected by small diffusion-time differences between M^AL^D and M^A; larger differences might change the probability of detecting M^AL^D vs. M^A, which can be accounted for by examining standards of pure M^AL^D and M^A.

Application of FAMS to Protein–DNA Interactions. We studied the sequence-specific interaction of *E. coli* CAP (16) with DNA as a model for protein–nucleic acid interactions. The diffusion time of CAP–DNA (570 μs) is comparable to that of free DNA (540 μs), allowing FAMS without detection-probability corrections. Specifically, we monitored the interaction of D-labeled CAP (as "ligand" CAP^D; Fig. 5*E*) with its consensus A-labeled DNA site (as "macromolecule" DNA^A), with or without allosteric effector cAMP.

With 0.2 mM cAMP, CAP^D–DNA^A complexes were detected as D–A species ($\theta \approx 0.50$, Fig. 5*C*; compare with DNA^A in Fig. 5*B*). CAP^D–DNA^A complexes appeared as a wide and heterogeneous *E* distribution, with a main peak (65% of all D–A species) at $E \approx 0.63$, a second peak (25%) at $E \approx 0.37$, and a shoulder (10%) at $E \approx 0.84$. High *E* values were consistent with a 60–100° overall DNA bend toward CAP (12, 17), whereas the heterogeneity, apart from reflecting the two possible D-labeling sites on CAP (Fig. 5*E*), might reflect heterogeneity caused by slow interconversion (slower than diffusion) between complexes with different magnitude of DNA bending. CAP^D–DNA^A complexes also showed a wide, slightly asymmetric *S* distribution (0.50 ± 0.15), possibly because of the small fraction of CAP^D with two donors (a D_2–A species, with *S* larger than for D–A species). Without cAMP, few complexes were formed ($\theta \approx 0.03$; Fig. 5*D*).

To assess equilibrium binding for the CAP–DNA interaction, we titrated 10 pM DNA^A with 0–300 pM active CAP^D, identified A-containing species, and calculated θ for each $[CAP^D]$ (Fig. 5 *A* and *F*). With cAMP, the dependence of θ to $[CAP^D]$ resembles a rectangular hyperbola (Eq. **5** with [D-only] = $[CAP^D]$); upon fitting, $K_d \approx 32 \pm 3$ pM, in good agreement with filter-binding-based values (24 ± 2 pM; ref. 18). Without cAMP, CAP binds to DNA >150-fold weaker ($K_d > 5$ nM).

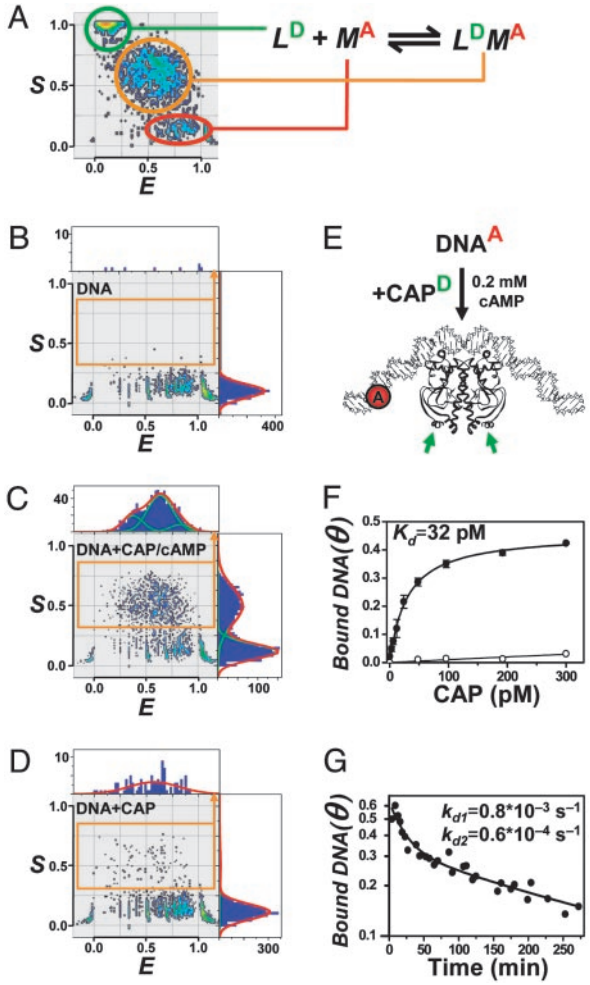

Fig. 5. Analysis of protein–DNA interactions using FAMS. (*A*) Analysis of interactions using FAMS. By using a labeled macromolecule (M^A; A-only species) and a labeled ligand (L^D; D-only species), we can monitor the formation of macromolecule-ligand complexes (M^AL^D; D–A species) on the *E-S* histogram by monitoring fractional occupancy θ at a ligand concentration $[L^D]$. Monitoring θ extracts equilibrium constants from histograms constructed at a single or multiple $[L^D]$, and kinetic constants from histograms constructed at multiple time points. (*B*) *E-S* histogram of A-containing species for 50 pM DNA^{A647}. (*C*) *E-S* histogram of A-containing species for 50 pM DNA^{A647}, 200 pM CAP^{TMR}, and 0.2 mM cAMP. (*D*) *E-S* histogram of A-containing species for 50 pM DNA^{A647} and 200 pM CAP^{TMR} (no cAMP). (*E*) Model of CAP–DNA complex and labeling scheme. The acceptor (A647, in red) was placed on DNA, and the donor (TMR) was placed on two possible sites on CAP (green arrows); because labeling efficiency was kept low, most D–A species have a single donor. (*F*) FAMS-based titration of DNA with CAP in the presence (filled circles) or absence of cAMP (open circles). With cAMP, CAP binds DNA with high affinity ($K_d \approx 32$ pM); without cAMP, CAP binds DNA weakly ($K_d > 5$ nM). (*G*) Kinetics of CAP–DNA dissociation, reflecting distinct rates of dissociation.

The fitted K_d for the CAP–DNA interaction (with cAMP) was expected to be slightly higher (≈20%) than the actual K_d, because we did not meet the requirement of $[DNA^A] \ll K_d$, which ensures that [free CAP^D] ≈ [total CAP^D] (15). However, [free CAP^D] can be determined directly from *E-S* histograms (Fig. 5*A*). Using the 200-pM CAP^D titration point as a standard to obtain the number of D-only bursts for 200 pM CAP^D, and the linearity between number of bursts and concentration (Fig. 11), we find [free CAP^D] ≈ 22 pM at 50% saturation (and therefore, $K_d \approx 22$ pM). The ability to measure [free L^D] allows probing of tight interactions (where it is difficult to satisfy $[M^A] \ll K_d$), and can account for inaccuracies caused by errors common when working with small volumes of liquids (e.g., losses on surfaces, pipetting errors).

When interactions involve single binding sites, and θ at saturation (θ_{sat}) is 1 or can be calculated (as in here, where labeling efficiency and ligand binding activities are known), we can combine a single histogram with Eq. **5** to calculate K_d for $[L^D]$ in the range of $(0.1–0.9) \times K_d$. The error is minimal when $[L^D] < K_d$ (at θ/θ_{sat} of 0.2, a 10% error in θ/θ_{sat} corresponds to 11% error in K_d), and increases at high θ/θ_{sat} (at θ/θ_{sat} of 0.8, a 10% error in θ/θ_{sat} corresponds to 50% error in K_d). Therefore, depending on the accuracy of θ/θ_{sat}, a 2D histogram at a single $[L^D]$ can yield an approximate K_d, or place it within limits.

We also monitored dissociation kinetics of CAP^D–DNA^A by forming the complex, diluting it in 50-fold molar excess unlabeled CAP (to sequester DNA^A formed because of dissociation), and observing the kinetics of θ decrease (Fig. 5*G*). If θ decrease is fitted as single exponential decay, dissociation rate constant $k_d \approx (1.1 \pm 0.2) \times 10^{-4}\ s^{-1}$, in agreement with gel-based assay values (ref. 19; $k_d \approx 1.2 \times 10^{-4}\ s^{-1}$). However, the θ decrease is clearly biexponential, with a fast ($k_{d1} \approx [1.2 \pm 0.8] \times 10^{-3}\ s^{-1}$) and a slow phase ($k_{d2} \approx [0.7 \pm 0.1] \times 10^{-4}\ s^{-1}$). This behavior has not been reported for CAP–DNA, but it has been reported for other protein–DNA complexes; it reflects an equilibrium between stable and unstable complexes, with stable complexes converting slowly to the unstable complexes, which dissociate (20, 21). It is possible that the small number of time points and partial dissociation of CAP–DNA during electrophoresis prevented observation of the fast phase by gel-based assays. FAMS can also monitor association kinetics, by fast mixing of low concentrations of CAP and DNA, and monitoring of the kinetics of θ increase.

Conclusion and Outlook

Using DNA fragments and a protein–DNA interaction, we showed that sorting of single molecules by using ALEX is an ultrasensitive method for simultaneous analysis of structure and interactions without separation steps; the interaction study also confirmed the quantitative nature of FAMS. FAMS can be readily used to study numerous biological processes that include simultaneous or sequential changes in distance and/or stoichiometry (Fig. 6), such as binding events that are followed by conformational changes (Fig. 6*A*), translocation events on linear tracks that culminate in dissociation of one of the components (Fig. 6*B*), assembly (or disassembly) pathways for large multicomponent complexes (Fig. 6*C*), and template-directed oligomerization on a nucleating template (Fig. 6*D*). It is important to note that analysis of such typical processes is either very complicated or impossible to perform with spFRET.

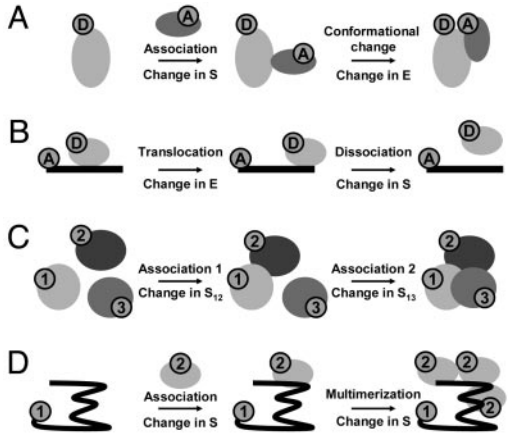

Fig. 6. Biological processes that can be dissected by simultaneous monitoring of structure and interactions. (*A*) Coupling of a binding event to a conformational transition. (*B*) Coupling of a translocation of a molecular motor and dissociation of a protein component. (*C*) Assembly of multicomponent complexes (using three fluorophores and additional stoichiometry-based emission ratios). (*D*) Template-directed multimerization (or aggregation) of biomolecules.

In a separate work, we have used FAMS to study the release of initiation factor σ^{70} from transcription complexes (A.N.K., R. Ebright, and S.W., unpublished data), to perform accurate FRET measurements on single-molecules (N.K.L., A.N.K., and S.W., unpublished data), and to study the kinetics and mechanism of abortive initiation of transcription. We have also used ALEX to study immobilized complexes (detected by using either confocal or total-internal-reflection microscopy) and deconvolve fluorophore photophysics from *E* measurements, thus permitting analysis of conformational dynamics.

Finally, valuable extensions of FAMS include expansion of the working concentration range to monitor low-affinity interactions [using excitation-volume confinement (22, 23), or nonfluorescent analytes that modulate interactions between fluorescent partners], use of additional excitations and fluorophores to monitor complex stoichiometries and multiple distances simultaneously, and combinations with fluorescence correlation spectroscopy or physical sorting (9).

We thank Drs. X. Michalet, L. Bentolila, and M. Jäger for the critical reading of the manuscript and discussions, Prof. R. H. Ebright for the generous gift of CAP and discussions, and J. Tang for editorial assistance. This work was funded by National Institutes of Health Grant GM65382:01 and by U.S. Department of Energy Grant DE-FG03-02ER63339 (to S.W.).

1. Clegg, R. M. (1992) *Methods Enzymol.* **211,** 353–388.
2. Ha, T., Enderle, T., Ogletree, D. F., Chemla, D. S., Selvin, P. R. & Weiss, S. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 6264–6268.
3. Deniz, A. A., Dahan, M., Grunwell, J. R., Ha, T., Faulhaber, A. E., Chemla, D. S., Weiss, S. & Schultz, P. G. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 3670–3675.
4. Ha, T., Rasnik, I., Cheng, W., Babcock, H. P., Gauss, G. H., Lohman, T. M. & Chu, S. (2002) *Nature* **419,** 638–641.
5. Schuler, B., Lipman, E. A. & Eaton, W. A. (2002) *Nature* **419,** 743–747.
6. Zhuang, X., Kim, H., Pereira, M. J., Babcock, H. P., Walter, N. G. & Chu, S. (2002) *Science* **296,** 1473–1476.
7. Dahan, M., Deniz, A. A., Ha, T., Chemla, D. S., Schultz, P. G. & Weiss, S. (1999) *Chem. Phys.* **247,** 85–106.
8. Ha, T., Ting, A. Y., Liang, J., Deniz, A. A., Chemla, D. S., Schultz, P. G. & Weiss, S. (1999) *Chem. Phys.* **247,** 107–118.
9. Eigen, M. & Rigler, R. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 5740–5747.
10. Li, H., Ying, L., Green, J. J., Balasubramanian, S. & Klenerman, D. (2003) *Anal. Chem.* **75,** 1664–1670.
11. Ren, X., Gavory, G., Li, H., Ying, L., Klenerman, D. & Balasubramanian, S. (2003) *Nucleic Acids Res.* **31,** 6509–6515.
12. Kapanidis, A. N., Ebright, Y. W., Ludescher, R. D., Chan, S. & Ebright, R. H. (2001) *J. Mol. Biol.* **312,** 453–468.
13. Laurence, T. A., Kapanidis, A. N., Kong, X., Chemla, D. S. & Weiss, S. (2004) *J. Phys. Chem. B* **108,** 3051–3067.
14. Mukhopadhyay, J., Kapanidis, A. N., Mekler, V., Kortkhonjia, E., Ebright, Y. W. & Ebright, R. H. (2001) *Cell* **106,** 453–463.
15. Gutfreund, H. (1995) in *Kinetics for the Life Sciences* (Cambridge Univ. Press, Cambridge, U.K.).
16. Busby, S. & Ebright, R. H. (1999) *J. Mol. Biol.* **293,** 199–213.
17. Schultz, S. C., Shields, G. C. & Steitz, T. A. (1991) *Science* **253,** 1001–1007.
18. Ebright, R. H., Ebright, Y. W. & Gunasekera, A. (1989) *Nucleic Acids Res.* **17,** 10295–305.
19. Fried, M. G. & Crothers, D. M. (1984) *J. Mol. Biol.* **172,** 263–282.
20. Dhavan, G. M., Crothers, D. M., Chance, M. R. & Brenowitz, M. (2002) *J. Mol. Biol.* **315,** 1027–1037.
21. Powell, R. M., Parkhurst, K. M., Brenowitz, M. & Parkhurst, L. J. (2001) *J. Biol. Chem.* **276,** 29782–29791.
22. Laurence, T. A. & Weiss, S. (2003) *Science* **299,** 667–668.
23. Levene, M. J., Korlach, J., Turner, S. W., Foquet, M., Craighead, H. G. & Webb, W. W. (2003) *Science* **299,** 682–686.