JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

ARTICLE

pubs.acs.org/JACS

Accurate Distance Determination of Nucleic Acids via Förster Resonance Energy Transfer: Implications of Dye Linker Length and Rigidity

Simon Sindbert,[†,∥] Stanislav Kalinin,[*,†,∥] Hien Nguyen,[‡] Andrea Kienzler,[§] Lilia Clima,[§] Willi Bannwarth,[§] Bettina Appel,[‡] Sabine Müller,[*,‡] and Claus A. M. Seidel[*,†]

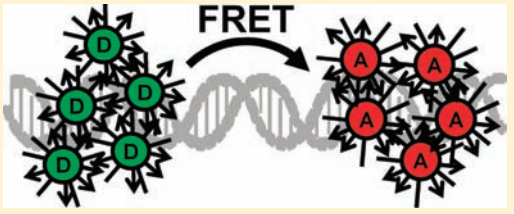
[†]Institut für Physikalische Chemie, Lehrstuhl für Molekulare Physikalische Chemie, Heinrich-Heine-Universität, Universitätsstrasse 1, Geb 26.32, 40225 Düsseldorf, Germany

[‡]Institut für Biochemie, Bioorganische Chemie, Ernst-Moritz-Arndt-Universität Greifswald, Felix-Hausdorff-Strasse 4, 17487, Greifswald, Germany

[§]Fakultät für Chemie und Biochemie, Albert-Ludwigs Universität Freiburg, AK Bannwarth, Albertstrasse 21, 79104, Freiburg, Germany

Supporting Information

ABSTRACT: In Förster resonance energy transfer (FRET) experiments, the donor (D) and acceptor (A) fluorophores are usually attached to the macromolecule of interest via long flexible linkers of up to 15 Å in length. This causes significant uncertainties in quantitative distance measurements and prevents experiments with short distances between the attachment points of the dyes due to possible dye−dye interactions. We present two approaches to overcome the above problems as demonstrated by FRET measurements for a series of dsDNA and dsRNA internally labeled with Alexa488 and Cy5 as D and A dye, respectively. First, we characterize the influence of linker length and flexibility on FRET for different dye linker types (long, intermediate, short) by analyzing fluorescence lifetime and anisotropy decays. For long linkers, we describe a straightforward procedure that allows for very high accuracy of FRET-based structure determination through proper consideration of the position distribution of the dye and of linker dynamics. The position distribution can be quickly calculated with geometric accessible volume (AV) simulations, provided that the local structure of RNA or DNA in the proximity of the dye is known and that the dye diffuses freely in the sterically allowed space. The AV approach provides results similar to molecular dynamics simulations (MD) and is fully consistent with experimental FRET data. In a benchmark study for ds A-RNA, an rmsd value of 1.3 Å is achieved. Considering the case of undefined dye environments or very short DA distances, we introduce short linkers with a propargyl or alkenyl unit for internal labeling of nucleic acids to minimize position uncertainties. Studies by ensemble time correlated single photon counting and single-molecule detection show that the nature of the linker strongly affects the radius of the dye's accessible volume (6−16 Å). For short propargyl linkers, heterogeneous dye environments are observed on the millisecond time scale. A detailed analysis of possible orientation effects (κ^2 problem) indicates that, for short linkers and unknown local environments, additional κ^2-related uncertainties are clearly outweighed by better defined dye positions.



1. INTRODUCTION

Measuring distances within biomolecules via Förster resonance energy transfer (FRET) has been a very useful technique in the field of structural biology for decades.[1−5] It is based on the fact that an excited fluorescent dye (donor) can transfer energy to another dye (acceptor) if the emission spectrum of the donor overlaps with the excitation spectrum of the acceptor. The efficiency of this energy transfer strongly depends on the distance between the dyes[3,6−9] allowing for donor−acceptor distance (R_{DA}) measurements in the range of about 20−100 Å. In the past years, FRET measurements on single molecules (smFRET) have become possible.[10,11] smFRET largely overcomes many problems of ensemble FRET, including species and time averaging, incomplete or unspecific labeling, as well as position-dependent donor or acceptor quenching artifacts.[12,13]

However, two problems commonly arise when performing quantitative FRET measurements. First, the fluorescent dyes are typically attached to the biomolecule via long flexible linkers, for example, the "standard" C6 (hexamethlyen) linker. The overall length of the linkage from the attachment point to the center of the chromophore is, thus, given by the length of the linker and the internal chemical structure of the dye and amounts to up to 20 Å. This yields a significant uncertainty in dye position and quenching environment.[14−18] Second, the FRET efficiency also depends on the relative orientation of the transition dipole moments of the two dyes,[3,7,8,19−21] which is expressed by the orientation factor κ^2. It can range from 0 to 4 and has a strong influence on the measured FRET efficiencies.

Received: June 29, 2010
Published: February 3, 2011

The effect of long dye linkers and the problems they can cause for quantitative FRET measurements are well-known.[2,22−24] There are approaches to circumvent this problem through modeling of the dye and linker motions,[25−29] and, thus, to calculate the accessible volume and mean position of the dye. This approach, however, works only for well-defined dye environments such as a straight double helix of double-stranded (ds) DNA or RNA. In many biomolecules, for example, complex RNA structures, the local structure is unknown, which makes such simulations impossible. Furthermore, there are approaches to minimize the dye position uncertainties, thus making modeling less important. Nucleobases can be replaced by fluorescent analogues,[30,31] and fluorophores can be covalently linked as end-caps of DNA duplexes[32] or stacked at the ends of the helices in the manner of additional base pairs.[21] However, this strongly restricts dye reorientation, and κ^2 cannot be assumed to be 2/3.

In this work, we suggest alternative procedures of obtaining highly accurate FRET-based structural information using internally labeled nucleic acids with dye linkers of different length and flexibility. Depending on whether the structure of the local environment of the dye is known, two cases have to be considered. If a simulation is possible, we demonstrate that an easily applicable accessible volume (AV) simulation method[28,29] provides realistic dye position distributions, which are consistent MD data and experimental smFRET results. For the case of an unknown local structure, we introduce alternative short dye linkers. They significantly decrease dye position uncertainty while allowing the fluorophores to rotate somewhat freely. In particular, short linkers are also expected to be more suitable for measuring short distances because, in this range, the length of long linkers becomes comparable to the absolute distances between donor and acceptor dye (R_{DA}). To consider the implications of using long and short dye linkers on quantitative FRET measurements, we present systematic studies for DNA and RNA to demonstrate the influence of the linker flexibility and length on the fluorescence properties of two representative dyes.

In section 2, we introduce the new short dye linkers for labeling of nucleic acids. We synthesized the modified nucleoside phosphoramidites shown in Figure 1 and incorporated them into oligonucleotides of defined sequence (section S1.1 in the Supporting Information). The alkenyl linkers (Figure 1A and B) were introduced by Heck chemistry,[33,34] and the propargyl linkers (Figure 1C and D) were introduced by Sonogashira coupling[33,35,36] starting from 5-iodo-2′-deoxyuridine or 5-iodoridine.

In section 3.1, we studied the influence of the linker flexibility and length on the individual fluorescence properties (fluorescence quantum yield, lifetime, and anisotropy) of Alexa488 and Cy5 dyes for DNA and RNA.

In sections 3.2 and 3.3, we characterized experimentally and theoretically the broadening of interdye distances due to dye linker motions. We performed quantitative FRET distance measurements using internally labeled dsDNA and dsRNA as test systems using single-molecule multiparameter fluorescence detection (smMFD) and ensemble time-correlated single-photon-counting (eTCSPC) techniques. We found that for both types of FRET measurements the modeling of the dye position and the R_{DA} distribution is essential for the quantitative interpretation of the observed FRET efficiencies. However, for short dye linkers, these corrections are much less important. For the simulation of the environment of fluorescent dye positions, we further developed simple geometric computations to calculate the sterically accessible volume (AV), which has been proposed to predict possible positions for EPR and FRET labels.[28,29] For defined environments, the AV approach provides dye position distributions that are very similar to those obtained by molecular dynamics (MD) simulations.[25] However, in contrast to the previous work,[29,37] we do not assume the dye to have a fixed static position within its AV but consider all accessible dye positions as equally populated. We demonstrate that this approach provides a far better approximation of the dye behavior. Furthermore, we account for the three different dimensions of the fluorophores. We applied this methodology in a FRET benchmark study for ds A-RNA with long C6 dye linkers, where a very good rmsd value of 1.3 Å was achieved. This demonstrates that long dye linkers can be safely used in FRET experiments, if three conditions are fulfilled: (1) the local structure of RNA or DNA in the proximity of the dye is known, (2) the DA distance is larger than the sum of the linkage lengths and larger than ~0.7 × Förster radius (that is, R_{DA} > 35−40 Å for our dyes), and (3) there are no stacking interactions between the dye and the nucleic acid. Using a multiparameter fluorescence detection setup[38] allows one to easily test for the presence of such interactions.

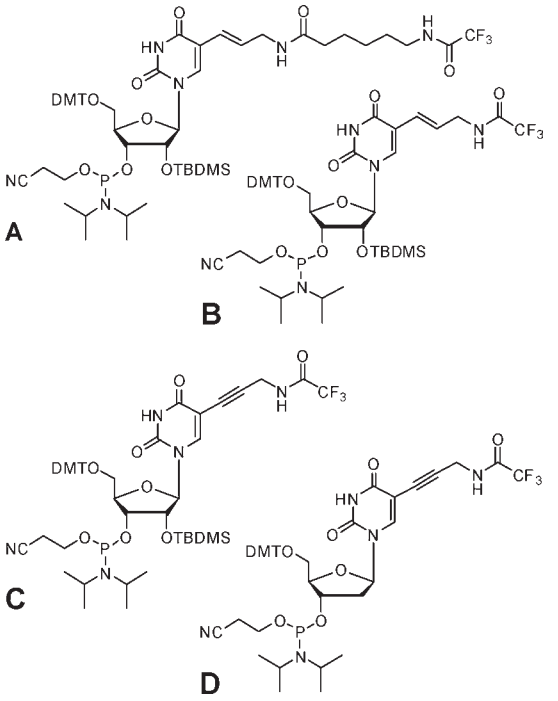


Figure 1. Modified RNA (A, B, and C) and DNA (D) building blocks with trifluoroacetamide-protected amino groups, which contain linkers of different length and flexibility.

In section 3.4, we studied the influence of the linker type on the additional broadening of R_{DA} distributions due to a very slow (>milliseconds) interchange between distinct dye environments. We can show that this effect becomes significant for short and stiff dye linkers.

In section 3.5, we introduce a rigorous procedure to minimize the uncertainties in the orientation factor κ^2, which is necessary because orientational distribution of both D and A is not strictly isotropic even for the longest linkers. We used time-resolved fluorescence anisotropy decays of D-only and A-only molecules and the FRET-sensitized acceptor anisotropy decay to determine the residual anisotropies, which allow us to compute probability distributions for possible values of κ^2 and estimate errors due to

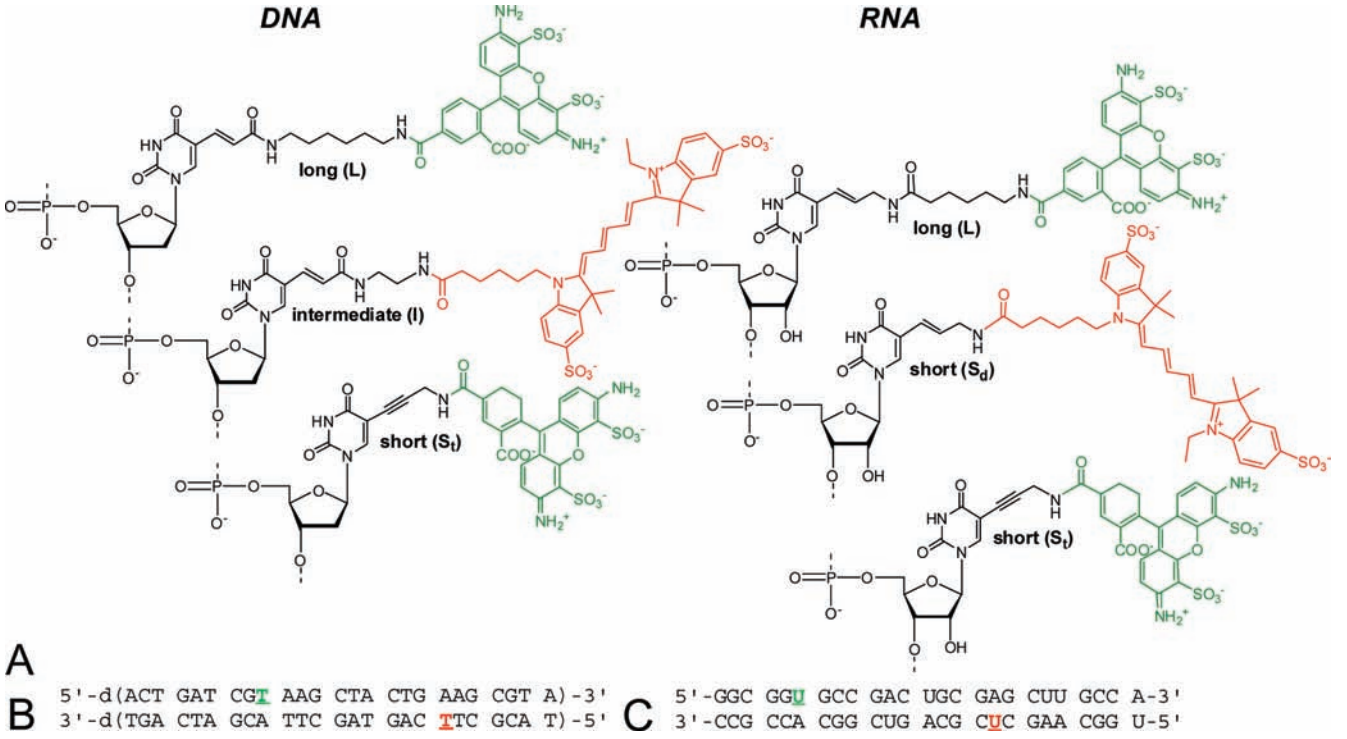


Figure 2. (A) Exemplary structures of the linkers S_t (propargyl) and S_d, I, and L (all alkenyl) with the dyes (Alexa488 and Cy5) used in DNA (left) and RNA (right). The donor and acceptor dyes are depicted in green and red color, respectively. Donor and acceptor dyes have been used in combination with each of the shown linkers. (B,C) Duplex structures of DNA1 (B) and RNA1 (C). Labeling positions are depicted in green and red for Alexa488 and Cy5, respectively.

uncertainties in the relative orientation of the dyes. Even if the residual anisotropies increase significantly for shorter linkers, it turns out that κ^2 errors increase only slightly. Thus, short and flexible linkers can be recommended for unknown local environments, because additional κ^2-related uncertainties are clearly outweighed by better defined dye positions.

2. MATERIALS AND METHODS

Synthesis of Labeled DNA and RNA. General. All reactions were carried out in dry solvents under argon atmosphere. All solvents and reagents were purchased from commercial sources and used as supplied unless otherwise stated. The solvents used in palladium coupling reactions were freed from oxygen. All products were visualized on TLC plates (aluminum sheets coated with silica gel 60 F 254, 0.2 mm thickness) at 254 nm ultraviolet light. Column chromatography was performed using silica gel type 60 ACC 35−70 μm. ^{1}H, ^{13}C, and ^{31}P NMR spectra were measured either on AC 250, ARX 300, AM 400, DRX 400, DRX 500, and DRX 600 systems from Bruker or on a Mercury VX 300 system from Varian using $CDCl_3$, CD_3CN, or [d_6]DMSO as solvent. UV spectra were recorded on a Perkin-Elmer-Lambda-35-UV/vis spectrometer or on a Varian Nano Drop ND-1000 spectrophotometer.

Linkers. The alkinyl linker, *N*-propargyltrifluoroacetamide (S_t), was synthesized according to Stockwell;[39] the alkenyl linkers, *N*-allyltrifluoroacetamide (S_d) and *N*-allyl-6-(*N*-trifluoroacetamido)hexanamide (L), were synthesized as described by Dey and Sheppard[36] and Meller and Brown.[40] See Figures 1 and 2 for the respective structures.

5-(3-Trifluoroacetamidopropargyl)-2′-deoxyuridine. A two-neck flask was charged with 5′-iodo-uridine (2.00 g, 5.65 mmol), Pd/C (304 mg, 0.282 mmol), CuI (214 mg, 1.13 mmol, 20 mol %), and Amberlite IRA 67 (5.14 g). The compounds were dried under high vacuum and afterward kept under argon. Propargyltrifluoroacetamide (1.72 g, 11.3 mmol, 2.3 equiv) and 56 mL of anhydrous DMF were added. Anhydrous argon was bubbled through this suspension for 5 min. The apparatus was again degassed and flooded with anhydrous argon. The reaction mixture was stirred at 50 °C for 17 h, and afterward cooled to room temperature and filtered over Celite. The Celite was washed with 50 mL of CH_2Cl_2/MeOH (5:1). The solvent of the combined organic phases was removed under reduced pressure. Remaining DMF was removed by condensation at 40 °C and 4 × 10^{-2} mbar. The crude product was purified by column chromatography on a Büchi Sepacore chromatography system (5 × 15 cm, F1−6, 20 mL; F7−30, 10 mL; $CHCl_3$/MeOH = 8.25:1.75). Fractions 7−30 were combined, and the solvent was removed under reduced pressure. The desired product was obtained as slightly brown foam. Yield: 79%. ^{1}H NMR (400 MHz, in [d_6]DMSO): δ = 2.08−2.14 (m, 2H, 2′-H), 3.51−3.63 (m, 2H, 5′-H), 3.78 (td, $^3J_{4',5'}$ = 3.4 Hz, $^3J_{4',3'}$ = 3.6 Hz, 1H, 4′-H), 4.20−4.22 (m, 3H, NH−CH_2 + 3′-H), 5.06 (t, $^3J_{5'OH,5'}$ = 5.1 Hz, 1H, 5′-OH), 5.22 (d, $^3J_{3'OH,3'}$ = 4.3 Hz, 1H, 3′-OH), 6.09 (t, $^3J_{1',2'}$ = 6.7 Hz, 1H, 1′-H), 8.17 (s, 1H, 6-H), 10.04 (m_c, 1H, N*H*−CH_2),11.6 (s, 1H, 3-NH). ^{13}C NMR (100 MHz, in [d_6]DMSO): δ = 161.5 (4-C), 156.0 (q, $^3J_{5'' 6''}F$ = 36.6 Hz, (C═O)CF_3), 149.4 (2-C), 144.1 (6-C), 115.8 (quart, $^2J_{6'' 6''}F$ = 288, CF_3), 97.6 (5-C), 87.6 (4′-C), 87.4 (1′′-C), 84.8 (1′-C), 75.4 (2′′-C), 70.2 (3′-C), 61.0 (5′-C), 40.2 (2′-C superposed by DMSO signals), 29.4 (3′′-C). MS (ESI): *m*/*z* (%) = 798.7 (13), 777.7 (11), 776.7 (37) [2M + Na^+], 701.9 (14), 550.9 (12), 399.8 (100) [M + Na^+].

5-(3-Trifluoroacetamidopropargyl)-5′-O-dimethoxytrityl-2′-deoxyuridine-3′-[(2-cyano-ethyl)-N,N′-diisopropylaminophosphoramidite] (Figure 1, **D**). 5-(3-Trifluoroacetamidopropargyl)-5′-*O*-dimethoxytrityl-2′-deoxyuridine was prepared according to the standard protocol for 5′-*O*-dimethoxytritylation of 2′-deoxynucleosides.[41] 5-(3-Trifluoroacetamidopropinyl)-5′-*O*-dimethoxytrityl-2′-deoxyuridine (430 mg, 0.6 mmol) and bis-diisopropylammonium tetrazolide (82 mg, 0.5 mmol) were dried three times azeotropically with 5 mL of anhydrous acetonitrile each. The remainder

dx.doi.org/10.1021/ja105725e |*J. Am. Chem. Soc.* 2011, 133, 2463–2480

was then dissolved in anhydrous CH_2Cl_2 (5 mL) and treated with 2-cyanoetoxy-bis-(*N*,*N*-diisopropylamino)phosphine (530 μL, 1.7 mmol). The reaction mixture was stirred for 3 h under argon at room temperature. Next, the mixture was poured into a degassed saturated $NaHCO_3$ solution and extracted three times with degassed CH_2Cl_2. The combined organic layers were dried over Na_2SO_4, and the solvent was removed under reduced pressure. The crude material was purified by short column chromatography on deactivated silica gel (1% Et_3N). The product was eluted with cyclohexane/AcOEt (5:7) and was obtained as a slightly beige foam. Yield: 83%. ^{1}H NMR (300 MHz, $CDCl_3$, both diastereomers): δ = 1.01−1.25 (m, 12H, 2 × CH($\underline{CH_3}$)$_2$,) 2.34−2.61 (m, 4H, CH_2CN, 2′-H), 3.32−3.82 (m, 12H, 5′-H OCH_2, 2 × *i*-Pr−CH, 2 × OCH_3), 3.87−3.91 (m, 1H, 4′-H), 4.14 (m, 2H, 9-H), 4.57−4.59 (m, 1H, 3′-H), 6.34 (t, ^3J = 6.6 Hz, 1H, 1′-H), 6.76−6.91 (m, 4H, ar.), 7.61−7.70 m, 9H, ar.), 8.21, 8.22 (2 × s, 1H, 6-H). ^{31}P NMR (120 MHz, $CDCl_3$) δ = 149.1, 150.0. MS (ESI): *m*/*z* (%) = 303.2 (14) [DMT^+], 901.9 (100) [M + 22$^+$].

5-(3-Trifluoroacetamidopropargyl)-5′-O-dimethoxytrityl-2′-O-tert-butyldimethylsilyl Uridine. 5-Iodoridine was 5′-*O*-tritylated and 2′-*O*-silylated according to standard procedures described in the literature.[42,43] 5-Iodo-5′-*O*-dimethoxytrityl-2′-*O*-*tert*-butyldimethylsilyl-uridine (300 mg, 0.38 mmol) was dissolved in 3.0 mL of DMF, and 228 μL of freshly distilled triethylamine was added. *N*-Propargyltrifluoroacetamide (182 mg, 1.1 mmol), $Pd(PPh_3)_4$ (44.0 mg, 0.066 mmol), and Cu(I)iodide (14.4 mg, 0.076 mmol) were added to the solution. The reaction mixture was stirred for 8 h in the dark at room temperature, concentrated in vacuo, and a solution of 5% Na_2EDTA was added to the residue. The crude product was extracted with ethyl acetate (3 × 20 mL). The combined organic layers were collected, dried over Na_2SO_4, and concentrated in vacuo. The crude product was purified by column chromatography (hexane/ethylacetate 70: 30) to give 5′-*O*-(4,4′-dimethoxytrityl)-2′-*O*-*tert*-butyldimethylsilyl-5-(3-trifluoroacetamidoprop-1-ynyl)uridine as a pale yellow solid. Yield: 66%. ^{1}H NMR (300 MHz, [d_6]DMSO, 25 °C, δ_H of the solvent at 2.5 ppm as internal reference): δ (ppm) = 0.03, 0.05 (2s, 6H, Si(CH_3)$_2$), 0.88 (s, 9H, C(CH_3)$_3$), 3.08, (2*s*, 2H, *H*-5′, *H*-5′′), 3.74 (s, 6H, 2 × OCH_3), 3.93 (d, ^3J = 5.4 Hz, 2H, N−CH_2), 4.01−4.09 (m, 2H, *H*-4′, *H*-3′), 4.34 (t, ^3J = 4.6 Hz, 1H, *H*-2′), 5.17 (d, ^3J = 6.0 Hz, 1H, O*H*-3′), 5.74 (d, ^3J = 5.7 Hz, 1H, *H*-1′), 6.89 (dd, ^2J = 9.0 Hz, ^3J = 3.3 Hz, 4H,*ar*), 7.23−7.43 (m, 9H,*ar*), 7.97 (s, 1H, *H*-6′), 9.97 (t, ^3J = 5.4 Hz, 1H, N−*H*), 11.74 (s, 1H, N−*H*). ^{13}C NMR (300 MHz, [d_6]DMSO, 25 °C, δ_C of the solvent at 39.61 ppm as internal reference): δ (ppm) = −5.14, −4.78, 17.92, 25.63, 29.31, 55.00, 59.75, 63.03, 69.62, 74.87, 75.56, 83.28, 85.92, 87.41, 88.86, 98.08, 113.22, 113.30, 115.76 (quart, ^2J = 289 Hz, CF_3), 126.65, 127.42, 127.93, 129.67, 129.73, 134.96, 135.54, 143.34, 144.84, 149.46, 155.93 (quart, ^3J = 37 Hz, (*C*═O)CF_3), 158.1, 161.41.

5-(3-Trifluoroacetamidopropargyl)-5′-O-dimethoxytrityl-2′-O-tert-butyldimethylsilyl-uridine-3′-[(2-cyanoethyl)-N,N′-diisopropylamino-phosphoramidite] (Figure 1, **C**). 5-(3-Trifluoroacetamidopropargyl)-5′-*O*-dimethoxytrityl-2′-*O*-*tert*-butyldimethylsilyl uridine (296 mg, 0.36 mmol) was coevaporated with 5 mL of dichloromethane containing 10% pyridine. The nucleoside was kept under vacuum overnight. Dry ethyl diisopropyl amine (0.32 mL, 4 × 0.36 mmol, freshly distilled just before use) was added, followed by 2 mL of dry dichloromethane. 2-Cyanoethyl-*N*,*N*′-diisopropylamino-chlorophosphoramidite (0.12 mL, 1.5 × 0.36 mmol) was added dropwise to the solution. After 3 h, 0.2 mL of dry methanol was added. The reaction mixture was diluted with 150 mL of ethyl acetate containing 10% triethylamine, washed with a saturated solution of Na_2CO_3 (1 × 10 mL), and with a saturated solution of potassium chloride (1 × 10 mL), dried over Na_2SO_4, and concentrated in vacuo. After purification by short column chromatography (hexane/ethyl acetate/triethyl amine 60: 30:10), the product was obtained as a pale yellow foam. Yield: 80%. ^{1}H NMR (300 MHz, [d_6]DMSO, 25 °C, δ_H of the solvent at 2.5 ppm as internal reference, both diastereomers): δ (ppm) = 0.01, 0.03, 0.05, 0.09 (4s, 12H, Si−CH_3), 0.83, 0.86 (2s, 18H, C(CH_3)$_3$), 1.18−1.53 (m, 12H, CH(CH_3)$_2$), 2.77 (t, ^3J = 5.8 Hz, 1H,CH_2−CN), 2.88 (t, ^3J = 5.8 Hz, 1H, CH_2−CN), 3.48−3.61 (m, 2H, C*H*(CH_3)$_2$), 3.74 (s, 6H, OCH_3), 3.97 (t, ^3J = 5.7 Hz, 2H, OCH_2), 4.16−4.22 (m, 2H, *H*-4′, *H*-3′), 4.49−4.56 (m, 1H, *H*-2′), 5.78 (d, ^3J = 5.7 Hz, 1H, *H*-1′), 5.84 (d, ^3J = 6.3 Hz, 1H, *H*-1′), 6.87−6.92 (m, 4H, *ar*), 7.27−7.43 (m, 9H, *ar*), 7.99, 8.01 (2s, 1H, *H*-6), 9.97−9.99 (m, 1H, N*H*), 11.75 (br s, 1H, N*H*). ^{31}P NMR (300 MHz, [d_6]DMSO, 25 °C, δ_P of H_3PO_4 at 0.0 ppm as the external reference): δ (ppm) = 148.24, 149.27.

5-(3-Trifluoroacetamidopropenyl)-uridine. 5-Iodoridine (740 mg, 2 mmol) was dissolved in 7 mL of DMF. To the resulting solution were added sodium acetate buffer (7.1 mL, 0.1 M, pH 5.2) and *N*-allyltrifluoroacetamide (2 mL, 17 mmol). A solution of $Na_2[PdCl_4]$ (658 mg, 2.2 mmol) in DMF (7 mL) was added while stirring vigorously. The reaction flask was placed in an oil bath at 80 °C for 8 h. The precipitated palladium was filtered off through Celite. The filtrate was concentrated in vacuo to a viscous brown oil. The crude product was purified by column chromatography (dichloromethane/methanol 98:2) to obtain a white powder. Yield: 58%. ^{1}H NMR (300 MHz, [d_6]DMSO, 25 °C, δ_H of the solvent at 2.5 ppm as internal reference): δ (ppm) = 3.55−3.70 (m, 2H, *H*-5′, *H*-5′′), 3.82−3.89 (m, 3H, CH_2, H-4′), 3.99 (1H, br s, *H*-3′), 4.07 (m, 1H, *H*-2′), 5.09 (br s, 1H, O*H*), 5.22 (br s, 1H, O*H*), 5.41 (d, 1H, O*H*), 5.77 (d, ^3J = 4.8 Hz, 1H, *H*-1′), 6.18 (d, ^2J = 15.9 Hz, 1H, ═C*H*), 6.46 (tt, ^2J = 15.9 Hz, ^3J = 6.1 Hz, 1H, ═C*H*), 8.12 (s, 1H, *H*-6), 9.69 (t, ^3J = 5.4 Hz, 1H, N*H*). ^{13}C NMR (300 MHz, [d_6]DMSO, 25 °C, δ_C of the solvent at 39.61 ppm as internal reference): δ (ppm) = 41.59, 60.51, 69.51, 73.73, 84.75, 88.13, 109.86, 115.97 (quart, ^2J = 288 Hz, CF_3), 123.92, 124.23, 138.20, 149.90, 156.06 (quart, ^3J = 36 Hz, (*C*═O)CF_3), 162.25.

5-[3-(6-Trifluoroacetylamidohexanamido)propenyl]uridine. 5-Iodoridine (740 mg, 2 mmol) was dissolved in 7 mL of DMF. NaOAc buffer (7.1 mL, 0.1 M, pH 5.2) and *N*-allyl-6-(*N*-trifluoroacetylamido)hexanamide (3.85 g, 14 mmol) were added to the solution, and a mixture of $Na_2[PdCl_4]$ (172 mg, 0.59 mmol) in DMF (2.5 mL) was added while stirring vigorously. The reaction flask was placed in an oil bath at 83 °C. After 2 h, another portion of $Na_2[PdCl_4]$ was added. After 8 h, the precipitated palladium was filtered off through Celite. $NaBH_4$ (2 × 12 mg) was added to the filtrate while vigorously stirring. The resulting yellowish solution was filtered through Celite, and the solvents were evaporated to give a viscous yellow oil. The crude product was purified by column chromatography, giving a white solid. Yield: 50%. ^{1}H NMR (300 MHz, [d_6]DMSO, 25 °C, δ_H of the solvent at 2.5 ppm as internal reference): δ (ppm) = 1.20−1.31 (m, 2H, CH_2), 1.42−1.55 (m, 4H, CH_2−CH_2), 2.09 (t, ^3J = 7.4 Hz, 2H, CH_2), 3.13−3.19 (m, 2H, CH_2), 3.54−3.60 (m, 2H, H5′, H5′′), 3.74 (t, ^3J = 5.4 Hz, 2H, CH_2), 3.82−3.87 (m, 1H, H4′), 3.99 (q, ^3J = 4.7 Hz, 1H, *H*-3′), 4.04−4.09 (m, 1H, *H*-2′), 5.07 (d, ^3J = 5.1 Hz, 1H, O*H*), 5.2 (t, ^3J = 4.8 Hz, 1H, O*H*), 5.39 (d, ^3J = 5.5 Hz, 1H, O*H*), 5.78 (d, ^3J = 4.9 Hz, 1H, *H*-1′), 6.12 (d, ^2J = 16.0 Hz, 1H, ═C*H*), 6.40 (tt, ^2J = 15.9 Hz, ^3J = 5.8, 1H, ═C*H*), 7.96 (t, ^3J = 5.6 Hz, 1H, N*H*), 8.08 (s, 1H, *H*-6), 9.39 (br s, 1H, N*H*), 11.42 (br s, 1H, N*H*). ^{13}C NMR (300 MHz, [d_6]DMSO, 25 °C, δ_C of the solvent at 39.61 ppm as internal reference): δ (ppm) = 25.11, 25.79, 28.04, 35.06, 40.80, 43.20, 60.55, 69.57, 73.68, 84.75, 88.04, 110.32, 115.95 (quart, ^2J = 289 Hz, CF_3), 122.09, 126.90, 137.34, 149.79, 156.10 (quart, ^3J = 36 Hz, (*C*═O)CF_3), 162.06, 171.69.

Both compounds, 5-(3-trifluoroacetamidopropenyl)uridine and 5-[3-(6-trifluoroacetyl-aminohexanamido)propenyl)uridine, were 5′-*O*-dimethoxytritylated and 2′-*O*-silylated with TBDMSCl according to standard protocols.[42,43]

5-(3-Trifluoroacetamidopropenyl-5′-O-dimethoxytrityl-2′-O-tert-butyldimethylsilyl-uridine-3′-[(2-cyanoethyl)-N,N′-diisopropylamin-ophosphoramidite] (Figure 1, **B**). 5-(3-Trifluoroacetamidopropenyl)-5′-*O*-dimethoxytrityl-2′-*O*-*tert*-butyldimethylsilyluridine (243 mg, 0.3 mmol) was coevaporated with dichloromethane (3 × 5 mL) containing 10% pyridine. The nucleoside was kept under vacuum overnight. Dry ethyl diisopropyl amine (0.27 mL, 1.2 mmol, freshly distilled over CaH_2 just before

used) was added, followed by 2 mL of dry dichloromethane. 2-Cyanoethyl-*N*,*N*′-diisopropylamino-chlorophosphoramidite (0.1 mL, 0.45 mmol) was added dropwise to the solution. After 3 h, 0.1 mL of dry methanol was added. After 15 min, the reaction mixture was diluted with 100 mL of ethyl acetate (prewashed with Na_2CO_3) containing 1% NEt_3, washed with a saturated solution of Na_2CO_3 (20 mL), dried over Na_2SO_4, and concentrated in vacuo to remove the solvents. The residue was purified by short column chromatography (hexane/ethyl acetate 60:40 → 50:50, 1% triethylamine) to get the product as a white foam. Yield: 87%. ^{1}H NMR (300 MHz, [d_6]DMSO, 25 °C, δ_H of the solvent at 2.5 ppm as internal reference, both diastereomers): δ (ppm) = 0.01, 0.03, 0.05, 0.08 (4*s*, 6H, Si−CH_3), 0.83, 0.85 (2*s*, 9H, C(CH_3)$_3$), 1.09−1.12 (m, 12H, CH(CH_3)$_2$), 2.77 (t, ^3J = 6.0 Hz, 2H, CH_2), 3.52−3.67 (m, 4H, H5′, H5″, NC*H*(CH$_3$)$_2$), 3.73 (s, 6H, OCH_3), 3.76−3.80 (m, 2H, OCH_2), 4.16−4.23 (m, 2H, *H*-4′, *H*-3′), 4.48−4.55 (m, 1H, *H*-2′), 5.65, 5.66 (2d, ^2J = 15.8 Hz, 1H, =C*H*), 5.84, 5.88 (2*d*, ^2J = 5.96 Hz, 6.38 Hz, 1H, *H*-1′), 6.25−6.36 (m, 1H, =C*H*), 6.87−6.91 (m, 4H, *ar*), 7.24−7.42 (m, 9H, *ar*), 7.68, 7.69 (2*s*, 1H, *H*-6), 9.56 (br s, 1H, N*H*), 11.59 (br s, 1H, N*H*). ^{31}P NMR (300 MHz, [d_6]DMSO, 25 °C, δ_P of H_3PO_4 at 0.0 ppm as the external reference): δ (ppm) = 148.27, 149.23.

5-[3-(6-Trifluoroacetylaminohexanamido)-propenyl]-5′-O-dimethoxytrityl-2′-O-tert-butyldimethylsilyluridine-3′-[(2-cyanoethyl)-N,N′-diisopropylaminophosphoramidite] (Figure 1, **A**). 5-[3-(6-Trifluoroacetylamidohexanamido)propenyl]-5′-*O*-dimethoxytrityl-2′-*O*-*tert*-butyldimethylsilyluridine (278 mg, 0.3 mmol) was coevaporated with dry pyridine (3 × 5 mL) and dichloromethane (3 × 5 mL). The nucleoside was kept under vacuum overnight. Dry ethyl diisopropyl amine (0.27 mL, 1.2 mmol, freshly distilled over CaH_2 just before used) was added, followed by 2 mL of dry dichloromethane. 2-Cyanoethyl-*N*,*N*′-diisopropylamino-chlorophosphoramidite (0.1 mL, 0.45 mmol) was added dropwise to the solution. After 2 h, another 0.1 equiv of phosphitylating reagent was added. After 2 h, 0.1 mL of dry methanol was added, and after 15 min the reaction mixture was diluted with 100 mL of ethyl acetate (prewashed with Na_2CO_3) containing 1% triethylamine, washed with a saturated solution of Na_2CO_3 (20 mL), dried over Na_2SO_4, and concentrated in vacuo. The residue was purified by short column chromatography (hexane/ethyl acetate 60:40 → 50:50, 1% triethylamine) to obtain a white foam. Yield: 82%. ^{1}H NMR (300 MHz, [d_6]DMSO, 25 °C, δ_H of the solvent at 2.5 ppm as internal reference, both diastereomers): δ (ppm) = 0.00, 0.02, 0.05, 0.08 (2*s*, 6H, Si(CH_3)$_2$), 0.83, 0.85 (2*s*, 9H, C(CH_3)$_3$), 1.11 (d, ^3J = 6.9 Hz, 12H, CH(CH_3)$_2$), 1.15−1.19 (m, 2H, CH_2), 1.43−1.48 (m, 2H, CH_2), 1.99−2.02 (m, 2H, CH_2), 2.75, 2.86 (2t, ^3J = 5.8 Hz, 2H, CH_2CN), 3.12−3.18 (m, 2H, CH_2), 3.24−3.30 (m, 2H, *H*-5′, *H*-5″), 3.46−3.59 (m, 2H, C*H*), 3.72−3.74 (m, 8H, OCH_3, CH_2), 3.77−3.84 (m, 2H, OCH_2), 4.14−4.16 (m, 1H, *H*-4′), 4.20−4.26 (m, 1H, *H*-3′), 4.47−4.57 (m, 1H, H-2′), 5.63, 5.65 (2d, ^2J = 15.8 Hz, 1H, =C*H*), 5.86, 5.90 (2d, ^3J = 6.4 Hz, 1H, *H*-1′), 6.17, 6.27 (m, 1H, =C*H*), 6.87−6.91 (m, 4H, *ar*), 7.23−7.43 (m, 9H, *ar*), 7.64, 7.67 (2*s*, 1H, *H*-6), 7.76−7.81 (m, 1H, N*H*), 9.39 (m, 1H, N*H*), 11.6 (br s, 1H, N*H*). ^{31}P NMR (300 MHz, [d_6]DMSO, 25 °C, δ_P of H_3PO_4 at 0.0 ppm as the external reference): δ (ppm) = 149.39, 148.22.

Oligonucleotides. Details on oligonucleotide synthesis[44,45] as well as on labeling of deoxyoligonucleotides and oligoribonucleotides with Cy5 and Alexa488 can be found in sections S1.1, S1.2, and S1.3 of the Supporting Information. Ultrapure labeled DNA1 with the L linker, RNA2 and RNA3 oligonucleotides (PAGE grade), and all unlabeled counter sequences were purchased from Purimex (Grebenstein, Germany). DNA1 oligonucleotides with the I Linker were purchased from IBA (Göttingen, Germany). All sequences of the DNA and RNA strands are given in section S1.4 of the Supporting Information. Additionally, the sequences and labeling positions of DNA1 and RNA1 are illustrated in Figure 2B and C. The linker types L and I for DNA and L for RNA were chosen for reasons of commercial availability, and the RNA linker types for reasons of chemical suitability. Throughout this work, if not stated differently, the same linker type is used at the donor (D) and the acceptor (A) positions.

Procedures . Hybridization of DNA and RNA. The hybridization buffer for the DNA samples contained 20 mM TRIS, 100 mM NaCl, and 10 mM $MgCl_2$, pH 7.5. For RNA, it contained 20 mM KH_2PO_4/K_2HPO_4, 100 mM KCl, and 10 mM $MgCl_2$, pH 6.5. The concentration of the DNA or RNA molecules in the buffer ranged between 2 and 10 μM. For FRET molecules, the ratio between the amount of acceptor and donor strand ranged from 1 to 2. For donor- and acceptor-only molecules, the ratio of unlabeled to labeled strand was 3:1. The solution was heated to a temperature of 90 °C inside a water bath and was then allowed to cool to room temperature overnight.

Time-Resolved Polarized Fluorescence Experiments and Data Analysis. Ensemble time-correlated single-photon-counting (eTCSPC) measurements were performed using pulsed laser excitation. Fluorescence intensity and anisotropy decay curves were fitted using the iterative reconvolution approach.[46] The fits approximately range from the maximum of the instrument response functions (IRF) to the first time channel with less than 100 detected photons. The fluorescence intensity decays of FRET-labeled molecules (donor and acceptor emission) were fitted globally with the decays of the molecules only labeled with either the donor (donor only, D-only) or the acceptor (acceptor only, A-only) dye. The fluorescence decays were modeled by single or double exponential decays or by assuming a Gaussian distribution of distances (section 3.3). Alternatively, the fluorescence decays were deconvoluted by using the maximum entropy method.[47,48] The anisotropy decays were recovered by globally fitting the sum ($F_\parallel + 2GF_\perp$) and difference ($F_\parallel - GF_\perp$) curves ($F_\parallel$, $F_\perp$, fluorescence signals in parallel and perpendicular polarization planes relative to the vertically polarized excitation light, respectively; G, ratio of the sensitivities of the detection system for vertically and horizontally polarized light). The anisotropy decays $r(t)$ were modeled by double or triple exponential decays (rotational correlation times ρ_1, ρ_2, and ρ_3) with free amplitudes (b_1, b_2, and b_3). For further details, see section S1.5 in the Supporting Information.

Single-Molecule Fluorescence Measurements. Multiparameter fluorescence detection (MFD) measurements were performed as described in refs 12,13,49. Each molecule generates a brief burst of fluorescence photons as it traverses the detection volume. This photon-train is divided initially into its parallel and perpendicular components via a polarizing beamsplitter and then into wavelength ranges using a dichroic beamsplitter. Bursts of fluorescence photons are distinguished from the background of 1−2 kHz by applying certain threshold intensity criteria.[50] For further details, see section S1.5 in the Supporting Information.

Measurements of Fluorescence Quantum Yields. Φ_F determination was performed according to ref 7. Rhodamine 700 in ethanol (Φ_F = 0.38)[51] and Rhodamine 110 (Φ_F = 0.95) were used as reference dyes for Cy5 and Alexa488, respectively. Correction for the solvent refractive index was performed as described in ref 7.

3. RESULTS AND DISCUSSION

For highly accurate FRET measurements, several linker and dye effects must be taken into account. First, the DNA and RNA microenvironment affects the dyes' photophysics and local motions. Second, all observable FRET parameters depend on spatial distributions of donor and acceptor positions. The final goal is to restrict the dye motions to achieve a well-defined dye localization in FRET experiments. If, on the other hand, the dye reorientation is restricted, additional orientational effects (κ^2 effects) finally influence the FRET efficiency, which we must learn to take into account.

3.1. Characterization of the Local Environment of the Dyes. As local quenching processes and restricted mobilities will complicate FRET analysis, we investigated how the nature and length of the linker influence the fluorescence properties of the dye for internally labeled dsDNA and dsRNA. Throughout

this work, we used Alexa488 as a donor and Cy5 as an acceptor dye (Figure 2). For internal postlabeling of the nucleic acids, we use the NHS-ester of Cy5, which, in contrast to phosphoamidite derivatives of Cy5 (having the same name, which leads to confusions; see Figure S2 in the Supporting Information), contains two sulfonic acid groups to prevent dye sticking. Three different linker types for the uracil or thymine base with decreasing stiffness and increasing length were used: (i) short stiff linkers with four backbone atoms (S_t and S_d); (ii) a linker of intermediate length (I) with seven backbone atoms; and (iii) long flexible linkers (L) with 11 backbone atoms, which are most frequently used in the scientific community (usually referred to as "C6-amino linker") (Figure 2).

3.1.1. Analysis of Local Quenching in DNA and RNA. We analyzed local quenching by fluorescence quantum yield and lifetime measurements. Figure 3A shows typical ensemble fluorescence lifetime measurements by eTCSPC (see Figure S1A and S1B in section S2.1 of the Supporting Information for the decays with the complementary FRET dye). We describe the fluorescence decays $F(t)$ of single-labeled dsDNA and dsRNA by up to two fluorescence lifetimes τ_i with the species fractions x_i and a species-averaged fluorescence lifetime $\langle\tau\rangle_x$ (eq 1).

$$F(t) = x_1 \exp(-t/\tau_1) + x_2 \exp(-t/\tau_2) \quad \text{with}$$
$$\langle\tau\rangle_x = x_1\tau_1 + x_2\tau_2 \qquad (1)$$

The results of the fluorescence lifetime analysis are summarized in Table 1. In most cases, the $\langle\tau\rangle_x$ values are proportional to the fluorescence quantum yields of the donor and the acceptor (Φ_{FD} and Φ_{FA}, respectively; see Table 1), which indicates purely dynamic quenching (see Table S1 in the Supporting Information).

The fluorescence of the rhodamine dye Alexa488 can be in principle quenched by nucleobases, which results in a multiexponential fluorescence decay. Photoinduced electron transfer (PET) or proton-coupled electron transfer between organic fluorophores and suitable electron-donating moieties, such as the nucleobase guanine, can quench fluorescence upon van der Waals contact.[15,52−54] PET quenching has been used as reporter for monitoring conformational dynamics in oligonucleotides.[54] It is striking that no quenching effects are observed for DNA, but there is quenching for RNA (see $\langle\tau\rangle_x$ in Table 1). This is in line with the spatial distribution of the dyes (see further). For RNA, the dye is closer to the nucleobases, because it is still partially inside the major groove, whereas in DNA it is primarily outside.

In contrast to Alexa488, Cy5 has its own additional linker with 6 atoms between the chromophore and the reactive coupling group (Figure 2), which increases the distance to the nucleobases, so that only slight fluorescence lifetime differences between DNA and RNA are noticeable.

The fluorescence properties of the Cyanine dye Cy5 are less affected by PET quenching[54] but rather more by trans → cis photoisomerization,[55−58] which is influenced by specific solvent effects[59,60] and sterical constraints set by the local environment.[58] In water, free Cy5 shows a single exponential fluorescence relaxation with a lifetime of 0.9 ns,[60,61] whereas our measurements yielded for each linker biexponential fluorescence decays of Cy5. The second lifetime was usually similar to that in water, and the first lifetime is significantly larger. In view of the above findings for the photoizomerization of Cyanine dyes, the multiexponential decay of Cy5 is most likely due to the heterogeneous DNA or RNA microenvironment and not necessarily due to significant

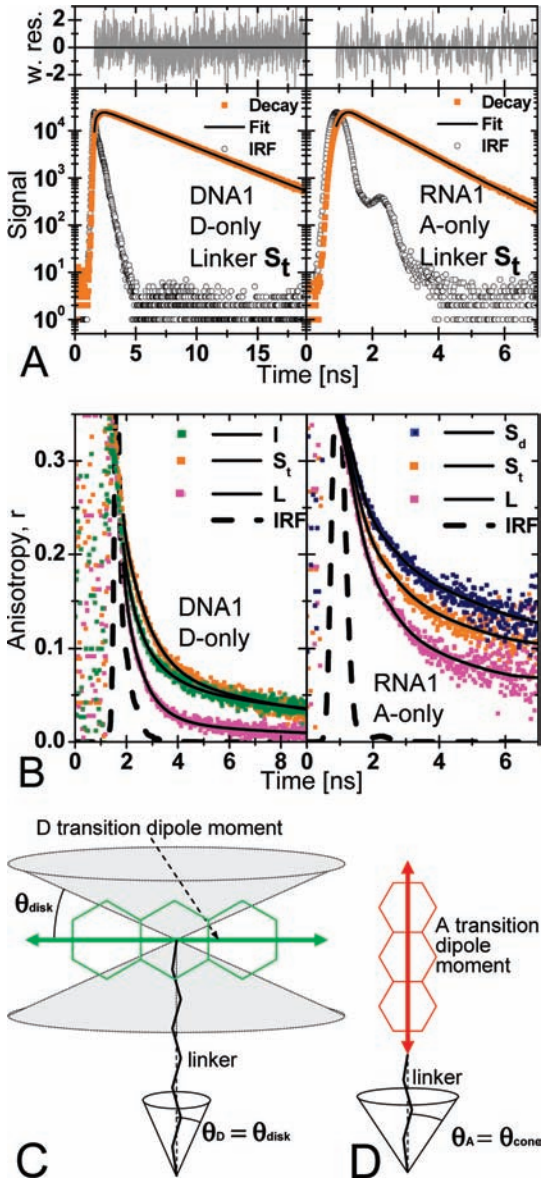


Figure 3. (A,B) Exemplary presentation of DNA1 (left panel) and RNA1 (right panel) eTCSPC measurements of single labeled nucleic acids: (A) fluorescence decay curves (weighted residuals are presented above each plot) and (B) fluorescence anisotropy decays with the rescaled IRFs curves. The fit results are listed in Tables 1 and 2. (C,D) Schematic sketch for the orientation of the transition dipole moments of the dyes with respect to the linker axes. The transition dipole moment of Alexa488 is assumed to be perpendicular to the linker axis; the linker can wobble within a cone with the opening half angle θ_D. (D) The transition dipole moment of Cy5 is assumed to be parallel to the linker axis; the linker can wobble within a cone with the opening half angle θ_A.

sticking of the dye to DNA or RNA. This interpretation of multiple microenvironments is supported by our recent single-molecule studies on Cy5-labeled dsDNA,[61] where at least two Cy5 states with distinct fluorescence lifetimes and anisotropies have been found. The steady-state anisotropy increases with the lifetime indicating a more restricted environment, which reduces the rate for cis−trans isomerization as well as the local linker wobbling motion (for more details, see section 3.1.2). Moreover, it is remarkable that for all linkers with close proximity to the allyl-unit (I for DNA

dx.doi.org/10.1021/ja105725e | J. Am. Chem. Soc. 2011, 133, 2463–2480

Table 1. Fluorescence Lifetimes (τ) and Quantum Yields (Φ_F) of Single Dyes Coupled to ds Nucleic Acids[a]

	D-only				A-only			
			DNA1					
linker	$\langle\tau\rangle_x$, ns	Φ_{FD}	τ, ns		$\langle\tau\rangle_x$, ns	Φ_{FA}	τ_1, ns (x_1)	τ_2, ns (x_2)
L	4.13	0.92	4.13		1.16	0.38	2.34 (11%)	1.01 (89%)
I	4.16	1.03	4.16		1.29	0.40	2.41 (16%)	1.08 (84%)
S_t	4.20	0.98	4.20		1.17	0.21	2.18 (12%)	1.03 (88%)
			RNA1					
linker	$\langle\tau\rangle_x$, ns	Φ_{FD}	τ_1, ns (x_1)	τ_2, ns (x_2)	$\langle\tau\rangle_x$, ns	Φ_{FA}	τ_1, ns (x_1)	τ_2, ns (x_2)
L	3.70	0.83	4.08 (90%)	0.27 (10%)	1.10	0.36	1.22 (73%)	0.79 (27%)
S_d	3.28	0.70	3.89 (81%)	0.67 (19%)	1.20	0.38	1.45 (51%)	0.93 (49%)
S_t	3.60	0.84	3.99 (88%)	0.71 (12%)	1.10	0.36	1.21 (77%)	0.72 (23%)

[a] Typical errors: single τ, ±0.02 ns; double τ_i (major component with $x_i \approx 80-90\%$), ±0.03 ns; τ_j (minor component with $x_j \approx 10-20\%$), ±0.4 ns; Φ_{FD}, ±0.05; Φ_{FA}, ±0.03.

and S_d for RNA) the mean fluorescence lifetimes $\langle\tau\rangle_x$ of Cy5 deviate from those of the other linkers, which indicates a slightly different mean environment.

Let us finally mention the only exception from the strict correlation between the species-averaged fluorescence lifetime $\langle\tau\rangle_x$ and the fluorescence quantum yield listed in Table 1 and Table S1. We surprisingly observe additional static Cy5 quenching for the S_t linker in DNA, which is also clearly detectable in ensemble- and sm-FRET experiments and will be discussed in detail in section 3.4.

3.1.2. Linker Motions in DNA and RNA. Figure 3B shows typical time-resolved ensemble measurements of fluorescence anisotropies (see also Figure S1C and S1D in section S2.1 of the Supporting Information for decays with the complementary FRET dye). The fluorescence anisotropy decays $r(t)$ with the fundamental anisotropy r_0 were formally characterized by up to three rotational correlation times ρ_i with the anisotropy amplitudes b_i (eq 2):

$$r(t) = b_1 \exp(-t/\rho_1) + b_2 \exp(-t/\rho_2) + b_3 \exp(-t/\rho_3) \quad \text{with} \quad r_0 \geq b_1 + b_2 + b_3 \qquad (2)$$

As the dye motion is partially restricted by the nucleic acids, the longest correlation time reflects to a significant extent the overall tumbling motion of the molecule (global motion; for more details, see section S1.5 of the Supporting Information), and its amplitude corresponds to the residual anisotropy r_∞, which allows the determination of second-rank order parameters $S^{(2)}$ (eqs 3 and 4). The average anisotropy corresponds to the steady-state anisotropy r_s, which is also measured in ensemble or single-molecule experiments by multiparameter fluorescence detection (MFD).

To rationalize dye motion, we must consider that the orientation of the transition dipole moment with respect to the linker is different for D and A. On the basis of the chemical structure of the linked donor dye Alexa488 in Figure 2, we assume the transition dipole moment to be approximately perpendicular to the linker axis as depicted in Figure 3C. Irrespective of the linker motions, rotations of the transition dipole about the linker axis significantly depolarize the donor fluorescence. In addition, if the linker can wobble within a cone with the opening half angle θ_D, the transition dipole of the donor can explore the space within a "disk" with the opening half angle $\theta_{disk} = \theta_D$ (Figure 3C). The second-rank order parameter $S_D^{(2)}$ is given by eq 3.

$$\text{donor}: \quad \frac{1}{2}\cos^2\theta_{disk} = \sqrt{\frac{r_{\infty,D}}{r_0}} = -S_D^{(2)} \qquad (3)$$

Thus, anisotropy senses both the linker and the dye rotations. Even if the linker cannot wobble (i.e., is totally stiff), the dye can still rotate about the linker, and a rather low residual anisotropy $r_{\infty,D} = 1/4(r_0)$ is expected.

In contrast, the transition dipole moment of the acceptor dye Cy5 is more parallel to the linker axis (Figure 2), and the linker together with the dye can wobble within a cone with the opening half angle $\theta_{cone} = \theta_A$ as depicted in Figure 3D. The motion is characterized by the second-rank order parameter $S_A^{(2)}$ in eq 4.[8]

$$\text{acceptor}: \quad \frac{1}{2}\cos\theta_{cone}(1+\cos\theta_{cone}) = \sqrt{\frac{r_{\infty,A}}{r_0}} = S_A^{(2)} \qquad (4)$$

Thus, anisotropy senses predominantly the linker motions as the dye rotates about the linker producing little or no fluorescence depolarization. If the linker cannot wobble, the dye rotates only parallel to the linker, which results in a very high residual anisotropy $r_{\infty,A} \approx r_0$. The approximation of the Cy5 motion by eq 4 is supported by the fact that $r_{\infty,A}$ is always >0.1 (Table 2), which is inconsistent with eq 3.

The analysis of the fluorescence anisotropies together with the rotational correlation times for D-only and A-only labeled DNA and RNA are compiled in Table 2.

As, in contrast to Cy5, the donor dye Alexa488 has no additional internal linker between the chromophore and the reactive coupling group, it is most suited to study the influence of the different nucleobase linkers. In RNA, the wobbling motion of the propargyl linker S_t and of the propenyl linker S_d is negligible (eq 3, $\theta_{disk} \approx 0°$), because the linker is stiff and short and the major groove of the RNA is very deep and narrow. The major groove of DNA is wider, and thus a small linker wobbling motion is observed ($\theta_{disk} = 16°$ and 23° for the S_t and I linkers, respectively). If the linkers become longer and more flexible, the linker wobbling should be limited not by the size of the dye but rather by the opening angle of the groove. Because Cy5 has its own

dx.doi.org/10.1021/ja105725e | J. Am. Chem. Soc. 2011, 133, 2463–2480

Table 2. Rotational Correlation Times ρ_i, Obtained for Donor-Only ($r_0 = 0.375$) and Acceptor-Only ($r_0 = 0.390$) DNA1 and RNA1 Samples[a]

	D-only					A-only			
linker	r_S	ρ_1, ns (b_1)	ρ_2, ns (b_2)	$\rho_{3(global)}$, ns ($b_3 = r_{\infty,D}$)[b]	θ_{disk}	r_S	ρ_1, ns (b_1)	$\rho_{2(global)}$, ns ($b_2 = r_{\infty,A}$)[b]	θ_{cone}
				DNA1					
L	0.043	0.17 (0.23)	0.76 (0.12)	7.9 [4−∞] (0.02)	45°	0.208	0.51 (0.26)	>60 [60−∞] (0.13)	47°
I	0.077	0.20 (0.21)	1.0 (0.10)	11.7 [8−25] (0.07)	23°	0.234	0.57 (0.22)	>60 [40−∞] (0.17)	41°
S_t	0.085	0.11 (0.13)	0.86 (0.16)	8.7 [6−11] (0.08)	16°	0.206	0.60 (0.25)	>60 [20−∞] (0.12)	47°
				RNA1					
L	0.065	0.27 (0.22)	1.5 (0.09)	6.6 [4−60] (0.05)	32°	0.211	0.68 (0.26)	8.1 [5−16] (0.13)	47°
S_d	0.128	0.38 (0.14)	2.2 (0.10)	9.5 [7−∞] (0.11)	∼0°	0.254	0.66 (0.18)	10.9 [8−15] (0.21)	35°
S_t	0.117	0.30 (0.15)	2.1 (0.09)	13.0 [8−15] (0.10)	∼0°	0.237	0.70 (0.21)	10.7 [8−19] (0.17)	40°

[a] Typical errors: r_s, ±0.002; ρ (major component), ±0.1 ns; r_∞, ±0.015. [b] For the longest (global) correlation time, 1σ confidence intervals are shown in squared brackets.

flexible 6 atom linker, the nucleobase linker effects are much weaker. Consistently, the L linker has the largest flexibility for both DNA and RNA. In DNA, the cone angles of both dyes, θ_{disk} and θ_{cone}, are equal, which supports the idea that the wobbling angle is only limited by the opening angle of the major groove. In RNA, the results still differ slightly because of the different dye linker lengths and the greater depth of the major groove. The experimental residual anisotropies nicely agree with the half opening angles of approximately 45° and 30° observed for the grooves of DNA and RNA, respectively (Table 2). As fluorescence lifetime measurements and single-molecule studies[61] indicate heterogeneous microenvironments of Cy5, it is important to note that the additional state with the longer lifetime is to some extent less mobile. The steady-state anisotropy r_s of this state is about 0.2,[61] indicating that the dye is not stuck but has sufficient rotational freedom (having in mind large values of $\rho_{(global)}$, $r_s \approx r_0$ would be expected for a immobile species).

To better understand the longest (global) rotational correlation time $\rho_{(global)}$ (Table 2), we performed simulations of DNA1 and RNA1 rotations using the HydroPro software.[62] For DNA1 and RNA1, respectively, three correlation times of 10, 22, and 35 ns (DNA1) and 12, 23, and 32 ns (RNA1) are predicted. In most cases, experimentally obtained values of $\rho_{(global)}$ (Table 2) are similar to the shortest predicted correlation time, which represents the rotation about the helical axis. A notable exception is Cy5 attached to DNA1, which shows systematically longer values of $\rho_{(global)}$. This fact suggests that the preferential orientation of the Cy5 transition dipole is nearly parallel to the helical axis of DNA, so that it senses the other slower rotational motions.

3.2. FRET Benchmark Study. In this and the following sections, we will show that it is crucial for a proper interpretation of FRET results to consider the distributions of D and A dye positions given by the vectors $\boldsymbol{R}_D$ and $\boldsymbol{R}_A$, respectively. Because different techniques determine distinct average distances between donor and acceptor dyes, we have to define and distinguish three different quantities.

(i) $\langle R_{DA} \rangle$ denotes the mean distance between the dyes and can be determined by eTCSPC measurements. $\langle R_{DA} \rangle$ is calculated by integrating over all possible positions of the two dyes and the resulting distances ($\langle R_{DA} \rangle = \langle |\boldsymbol{R}_D - \boldsymbol{R}_A| \rangle$).

(ii) $\langle R_{DA} \rangle_E$ is the FRET-averaged distance between the dyes. It is calculated from the mean FRET efficiency (eq 5A) using eq 5B.

$$\langle E \rangle = \left\langle \frac{1}{1 + R_{DA}^6 / R_0^6} \right\rangle \quad (5A)$$

$$\langle R_{DA} \rangle_E = R_0 (\langle E \rangle^{-1} - 1)^{1/6} \quad (5B)$$

In eq 5, R_0 is the Förster radius. $\langle R_{DA} \rangle_E$ is determined from time-averaged fluorescence intensity measurements on the single-molecule (section 3.2.2) or ensemble level.

(iii) R_{mp} is the distance between the mean positions of the dyes ($R_{mp} = |\langle \boldsymbol{R}_D \rangle - \langle \boldsymbol{R}_A \rangle|$) and is used for the geometric description of FRET-based structural models.[25] As shown below, R_{mp} cannot be measured directly via FRET. The detailed calculation of $\langle R_{DA} \rangle$, $\langle R_{DA} \rangle_E$, and R_{mp} is described in section S2.2 (eqs S3−S5) in the Supporting Information.

3.2.1. Calculation of the Volume Accessible to the FRET Dyes

The AV Approach. The prediction of the FRET dye positions with respect to the macromolecule of interest is absolutely essential for the interpretation of quantitative FRET measurements, especially when the dyes are attached via long flexible linkers. If the local structure of a macromolecule is known or can be predicted, the dye positions have been successfully computed by molecular dynamics (MD) simulations.[25−27] However, MD simulations are time-consuming and too complex for everyday use. Recently, alternative methods based on simple geometric computations have been proposed to predict possible positions for EPR and FRET labels.[28,29]

As sketched in Figure 4A, these methods approximate the dye by a sphere with an empirical radius of R_{dye}, where the central atom of the fluorophore (see Table 3 for definition) is connected by a flexible linkage of a certain effective length L_{link} and width w_{link} to the nucleobase. The overall length of the linkage is given by the actual length of the linker and the internal chemical structure of the dye. A geometric search algorithm finds all dye positions within the linkage length from the attachment point, which do not cause steric clashes with the macromolecular surface. All allowed positions are considered as equally probable, which allows one to define an accessible volume for the dye (AV).

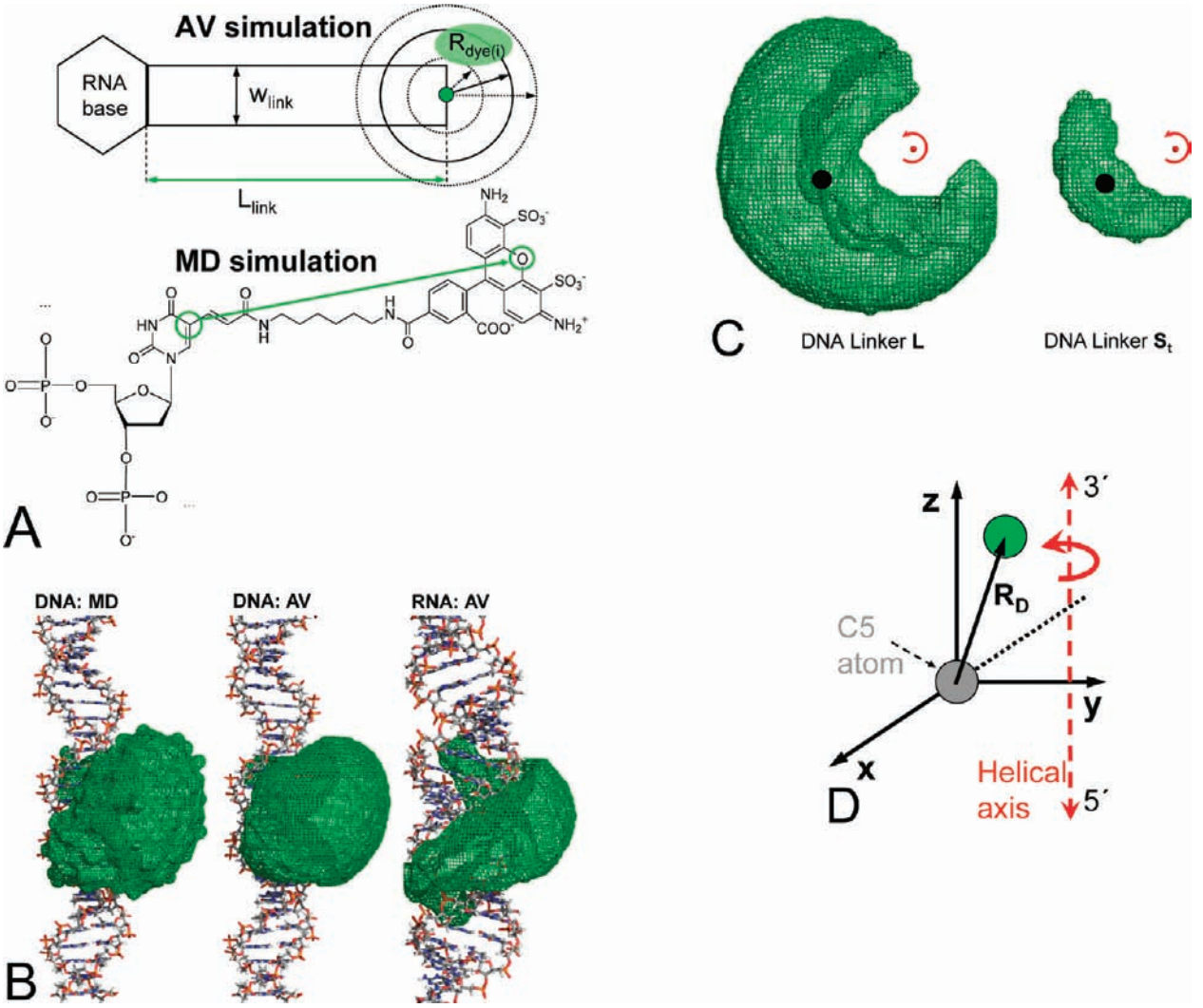


Figure 4. (A) Schematic comparison of AV and MD simulations of dye positions. For AV simulations, we used w_{link} = 4.5 Å, and $R_{dye(1)}$ = 5 Å, $R_{dye(2)}$ = 4.5 Å, and $R_{dye(3)}$ = 1.5 Å for the three dimensions of Alexa488. See Table 3 for values for L_{link}. (B) Molecular dynamics (left) and AV (middle, right) simulations of possible positions of Alexa488 attached to DNA1 (B-form) (left, middle) or RNA1 (A-form) (right) via linker L. The structures are rendered via PyMOL.[65] (C) Top view of Alexa488 position distributions simulated by the AV approach for L (left) and S_t (right) linkers. The helical axis of DNA and the mean position of the dye are shown as red and black spheres, respectively. (D) Coordinate system used to define the mean dye positions $\langle \boldsymbol{R}_D \rangle$ and $\langle \boldsymbol{R}_A \rangle$ in Table 3. The red dashed arrow indicates the helical axis of the nucleic acid.

Spatial Requirements Are Better Described by a Complex Fluorophore Shape. For RNA, AV simulations using the empirical dye radius R_{dye} = 3.5 Å[29] result in two separated dye clouds (see Figure S3 in the Supporting Information). It is obvious that the space between the obtained clouds should also be accessible for the planar chromophores used here. This demonstrates that, for some sterically demanding local environments, for example, in RNA, it is important to take the three quite different dimensions of a fluorophore into account.

Therefore, for each calculation of a position distribution, we used the real physical dimensions of the fluorophore and performed three independent AV simulations with three different radii $R_{dye(i)}$ and superimposed them. Thus, the obtained position distribution represents an average weighted by the number of allowed positions. Throughout this work, we used for Alexa488 $R_{dye(1)}$ = 5 Å, $R_{dye(2)}$ = 4.5 Å, and $R_{dye(3)}$ = 1.5 Å and for Cy5 $R_{dye(1)}$ = 11 Å, $R_{dye(2)}$ = 3 Å, and $R_{dye(3)}$ = 1.5 Å. It turned out that these "mixed" AV simulations are necessary to accurately predict dye distributions for RNA; yet for DNA, the effect of the dye radius is much less pronounced.

The fact that the dyes are assumed as spheres makes it impossible to take into account aberrations due to asymmetric structures (e.g., for Cy5, Figure 2). However, in our case, this results only in an angular shift, which is easy to correct for as we have the possibility of comparison to MD data.

The AV method is clearly not applicable when dyes show considerable interactions (such as sticking) with DNA or RNA.[21] However, in this work, we employ internal labeling of DNA and RNA, which minimizes interactions of the dyes with DNA and RNA. We use NHS-ester of Cy5, which, in contrast to phosphoamidite derivatives of Cy5 (see Figure S2 in the Supporting Information), contains two sulfonic acid groups. The negative charges of Cy5 and Alexa488 largely prevent dye sticking. With the exception of dyes with S_t linkers (see section 3.4), there is no evidence for the presence of long-lived dye heterogeneities, which justifies the use of the AV method in this work. It is worth

dx.doi.org/10.1021/ja105725e |*J. Am. Chem. Soc.* 2011, 133, 2463–2480

Table 3. Estimation of Mean Positions for Alexa488 and Cy5 Using MD and AV Simulations

method	linker	L_{link}, Å (D)	$\langle R_D \rangle$, Å (x, y, z)[a]			$\lvert\langle R_D \rangle\rvert$, Å[a]	σ_D, Å	L_{link}, Å (A)	$\langle R_A \rangle$, Å (x, y, z)[a]			$\lvert\langle R_A \rangle\rvert$, Å[a]	σ_A, Å
						DNA1							
MD	L		7.4	−3.8	4.4	9.4	11.4		5.0	−7.1	2.7	9.1	12.9
AV[b]	L	20	6.9	−4.1	2.2	8.3	13.1	22	6.5	−4.0	2.3	7.9	15.2
	I	15	6.0	−4.3	3.1	8.0	9.0	17	6.2	−4.0	2.7	7.9	11.0
	S_t	11	4.3	−4.7	3.2	7.1	5.7	16	6.1	−4.0	2.9	7.8	10.2
						RNA1							
AV[b]	L	20	−0.1	−8.5	4.5	9.6	11.3	22	−0.8	−8.5	4.2	9.5	12.8
	S_d	11	−4.0	−2.9	3.9	6.3	4.9	14	−3.9	−3.4	3.7	6.3	6.9
	S_t	11	−4.0	−2.9	3.9	6.3	4.9	16	−2.8	−4.9	3.9	6.9	8.6

[a] Between the C5-atom of the base (origin) and for Alexa488 the O atom at position 10 of the xanthene ring or for Cy5 the C atom at position 3 of the pentamethine chain; see Figure 4D. [b] For all AV simulations, we used the same linkage width w_{link} = 4.5 Å. The distinct linkage lengths L_{link} were determined from the most extended conformations.

mentioning that strong interactions between a dye and DNA or RNA are probably impossible to adequately model even by MD simulations, because millisecond time scales are currently not accessible to MD.[63] In this work, we developed an improved accessible volume simulation procedure based on the algorithm "Model Satellite Prior" implemented in the "FRETnpsTools" program[29,37] and then performed our own distance calculations as described. The pdb files of the macromolecules were generated with the Nucleic Acid Builder (NAB) software, which is part of AmberTools.[64]

Comparison of AV with MD. We tested the suitability of the AV approach in two steps: (i) in this section, the predictions of the AV approach are compared to the results of the MD simulation from;[25] and (ii) in section 3.2.2, the AV approach is used to model the dye position distributions in a FRET benchmark study.

Figure 4B shows distributions of possible positions of Alexa488 attached to DNA and RNA via the L linker simulated by MD (left) and AV (middle and right). Each AV simulation needs five input parameters: three dye radii $R_{dye(1,2,3)}$ as defined above, w_{link}, and L_{link}. We used typical parameters for the linkage width (w_{link} = 4.5 Å) from ref 29. The linkage lengths (L_{link}) were estimated from the fully extended conformations of each linker using the Hyperchem software[66] and are listed in Table 3.

Considering the L linker, the outer border of the volume accessible to Alexa488 attached to DNA is displayed as a green net in Figure 4B. The volume calculated by AV closely resembles the distribution predicted by MD. In comparison to MD data, AV predicts the mean position (of Alexa488 ($\langle R_D \rangle$)) (O atom at position 10 in the xanthene ring) with respect to the C5 atom of the uracil (Figure 4D) with a deviation of 2.2 Å (Table 3). As expected, due to the asymmetric structure of Cy5, the distribution of its positions simulated using the AV approach agrees less well with MD data (3.5 Å deviation between the respective mean positions $\langle R_A \rangle$ defined by the C atom at position 3 in the pentamethine chain). However, the z-displacement (2.3 Å; see Figure 4D) and the distance from $\langle R_A \rangle$ to the helical axis (11.5 Å) are similar to MD values (2.7 and 11.8 Å, respectively). Therefore, the main difference between the mean positions of Cy5 predicted by MD and AV is a small angular displacement along the xy-plane of ~20°. Thus, in the following, we apply this additional shift for all Cy5 positions predicted by AV.

Regardless of the linker and the fluorophore, the z-coordinates of the mean positions of the dyes with respect to the C5 atom of the nucleobase ($\langle R_D \rangle$ and $\langle R_A \rangle$) are always positive (Figure 4D). This means that they are always shifted toward the 3′-end of DNA or RNA, which has a significant effect on DA distances (inset in Figure 5C). For L linkers, the displacement between the dyes in z-direction due to linkers (Δz_{link}) is 7−8 Å (Table 3). Thus, for structure determinations or when choosing labeling positions, it is crucial to take this displacement into account. Additionally, as opposed to DNA, the base pair plane in A-RNA is not perpendicular to the helical axis so that the C5 atoms of opposing nucleobases are displaced against each other by Δz_{C5} = 1.6 Å toward the 3′-ends of the respective strands. In summary, the z-displacement of the dyes is determined by three factors and is given by $\Delta z = \Delta z_{bp}\Delta n + \Delta z_{link} + \Delta z_{C5}$ where the basepair separation Δn is counted from D to A toward the 5′-end of DNA or RNA (i.e., Δn is negative if A is closer to the 3′ end).

For both DNA and RNA, the attached dyes point into the major groove, and the volume follows the helical twist of the groove, which is visible best for RNA (Figure 4B). The groove is deeper and narrower for RNA, so that the dye motion is expected to be more restricted. This is consistent with the experimental anisotropy decays (Section 3.1.2).

In the case of DNA, the calculated mean dye positions are remarkably insensitive to linkage length (Table 3), which is probably due to the fact that, with increasing linker length, the accessible volume expands in all directions as illustrated in Figure 4C. On the other hand, the distance distribution shape changes significantly with linkage length used for the simulations (Figure 4C). As discussed below, all observable FRET parameters can strongly depend on the width of the R_{DA} distribution. In other words, the knowledge of the mean dye positions is usually insufficient to predict the FRET efficiency. Thus, and to compare AV with MD results, the standard deviations of the simulated position distributions (σ_D and σ_A) are also included in Table 3. For the L linkers, the height (projection to the z-axis, Figure 4D) of the accessible volume corresponds to ±5 base pairs, which results in significant problems for short DA distances.

3.2.2. Single-Molecule FRET Measurements of RNA Prove the Accuracy of the AV Model. For RNA (Figure 4B, right), no MD data are available to calibrate the AV parameters. To test the predictions of the AV approach nevertheless, we experimentally determined five DA distances within two internally labeled dsRNAs (RNA2 and RNA3, sequences and labeling positions can be found in section S1.4 in the Supporting Information) via

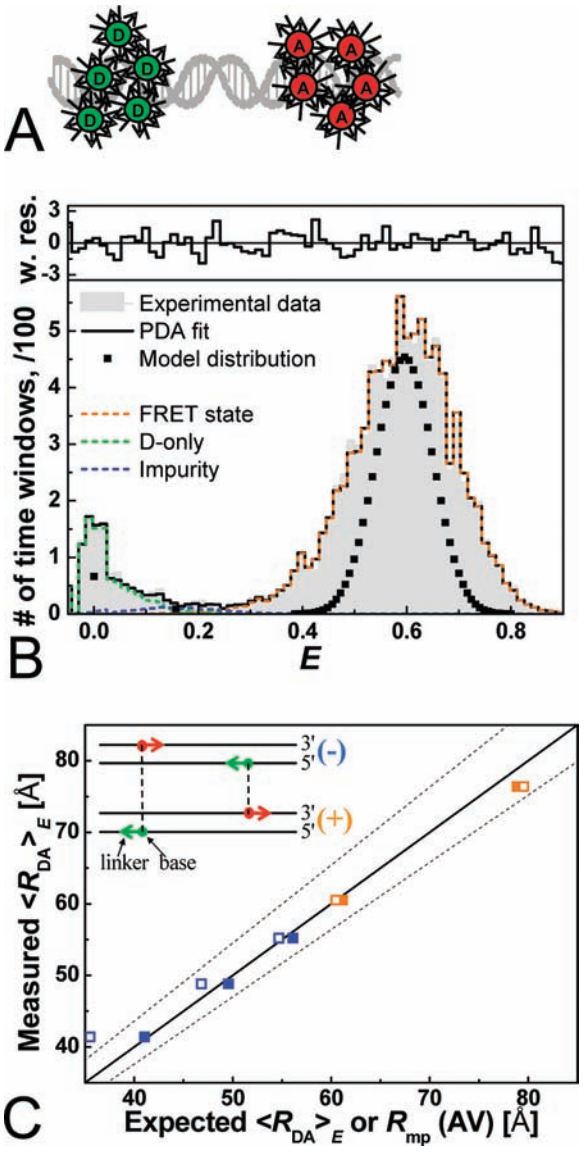


Figure 5. (A) Sketch illustrating the dynamic orientation averaging of FRET: the diffusion of the dyes in the accessible volume with the characteristic diffusion rate constant k_d and the reorientation fluctuations with the rate k_R ($k_R = 1/\rho_1$); $k_R \gg k_{FT} \gg k_d$ (k_{FT}: FRET rate constant). (B) PDA of RNA2(19−) (selected bursts). FRET efficiency histogram of experimental data (gray area) is fitted (black solid line) using the following parameters: $\langle R_{DA}\rangle_E$ = 48.8 Å; σ_{app} = 1.9 Å; 12.5% of D-only; 1.6% of impurities with apparent R_{DA} = 68.8 Å (also present in D-only samples); $\chi^2_r = 0.91$. Weighted residuals are shown in the upper plot. PDA parameters: time window Δt = 1 ms; mean background intensities in the green and red detection channels $\langle B_G\rangle$ = 1.23 kHz; $\langle B_R\rangle$ = 0.49 kHz; spectral crosstalk, 1.7%; $\Phi_{FD(0)}$ = 0.8; Φ_{FA} = 0.29; Alexa488−Cy5 Förster radius R_0 = 52 Å; green/red detection efficiency ratio: 0.78. (C) DA distances $\langle R_{DA}\rangle_E$ measured by smFRET (compiled also in Table S2 in section S2.5 of the Supporting Information) plotted versus expected (simulated using AV) distances $\langle R_{DA}\rangle_E$ (■) and R_{mp} (□). The following polynomial approximation of a R_{mp}↔$\langle R_{DA}\rangle_E$-conversion function was used: $\langle R_{DA}\rangle_E = -2.68 \times 10^{-5} R_{mp}^3 + 7.53 \times 10^{-3} R_{mp}^2 + 0.272 R_{mp} + 23.1$. The solid line represents equal expected and measured distances. The statistical experimental errors are smaller than the symbol size. The dashed lines represent the expected uncertainties due to possible errors of κ^2 (section 3.5) given by the typical precision of 5.1% in Table 5. Inset: Sketch illustrating that for a given basepair separation between D and A, the linker orientation may lead to a decrease (−, blue) or an increase (+, orange) in R_{DA}.

smFRET. Moreover, we used these two sequences to demonstrate also the effect of a dye displacement toward the 3′ end of the nucleic acid caused by the linker (inset of Figure 5C). In RNA2, both dyes are close to the 5′ ends, so that they are displaced toward each other (− effect). In RNA3, both dyes are close to the 3′ ends and are thus displaced in opposite directions (+ effect). To avoid any possible orientational artifacts in this benchmark study, we investigated RNAs, where the dyes were attached by L linkers.

From FRET Data $\langle R_{DA}\rangle_E$ to Structural Information R_{mp}. As we will show in sections 3.3 and 3.4, distinct time averaging regimes, from nano- to milliseconds, need to be considered for the interpretation of FRET data acquired with different experimental techniques. For this section, it is sufficient to know that all fluorophore positions of molecules with L linkers will be averaged during the millisecond dwell time of the single molecule in the confocal observation volume. Thus, the information on the width of the position distribution is lost; that is, the mean FRET efficiency $\langle E\rangle$ is observed. Because FRET efficiencies and distances are averaged differently, that is, $E(|\langle \boldsymbol{R}_D\rangle - \langle \boldsymbol{R}_A\rangle|) \neq \langle E(|\boldsymbol{R}_D - \boldsymbol{R}_A|)\rangle$, simulated mean DA position distances (R_{mp}) cannot be directly compared to the experimental DA distance values $\langle R_{DA}\rangle_E$ (eq 5). As $\langle R_{DA}\rangle_E \neq R_{mp}$, one cannot measure R_{mp} directly. However, for structure determination, it is necessary to obtain distances that allow a comparison between measurement and simulation. For solving this problem, several features must be considered. The easiest solution (algorithm 1) is to calculate R_{mp} from measured $\langle R_{DA}\rangle_E$ or E by applying a R_{mp}↔$\langle R_{DA}\rangle_E$ conversion function (for details, see ref 25). However, especially for small distances, this function can be ambiguous because the slope becomes very small (Figure S4 in the Supporting Information). Moreover, $\langle R_{DA}\rangle_E$ depends not only on R_{mp} but also to some extent on the mutual orientation of dye clouds. Thus, for a safe solution (algorithm 2), it is generally advantageous to calculate the theoretical $\langle R_{DA}\rangle_E$ values and directly compare with experimental data. Because the R_{DA} distribution is directly obtained from AV simulations, theoretical $\langle R_{DA}\rangle_E$ can be calculated according to eq 5. On the other hand, due to its high speed, the easy approach with the R_{mp}↔$\langle R_{DA}\rangle_E$ conversion function is more practical for structure modeling. For instance, if an iterative algorithm is applied to structure optimization, remodeling of the dyes' AVs and calculating $\langle R_{DA}\rangle_E$ at each iteration step may become very time-consuming, so that this cannot be done in all steps.

In eq 5A, two assumptions are made, which are sketched in Figure 5A and referred to as dynamic orientation averaging of FRET. First, the diffusion of the dyes in the sterically allowed volume with the characteristic diffusion rate constant k_d is much slower than the FRET rate constant k_{FT}. Here, by FRET rate, we mean the formal overall kinetic rate of donor quenching via FRET (typically some ns^{-1}), which can be directly measured using TCSPC. In other words, the distribution of individual DA distances $R_{DA,i}$ is quasi-static on the FRET time scale. Second, the local reorientation fluctuations with the rate constant of $k_R = 1/\rho_1$ are fast so that a mean effective orientation factor κ^2 (due to the restriction of dye motions it is not necessarily equal to 2/3) can be used. In this work, we will experimentally check the validity and error limits of assuming the "standard" case, that is, isotropic average with fast and unrestricted dye rotation ($k_R \gg k_{FT} \gg k_d$ and the wobble half angle 90°, resulting in $\kappa^2 = 2/3$).

Photon Distribution Analysis (PDA). When photon bursts of freely diffusing molecules are analyzed, it is mandatory to take into account that the obtained FRET efficiency histograms are affected by the stochastic nature of photon emission and detection

(shot noise) and other sources of dynamic or static heterogeneities. The use of an exact description for the theoretical shot noise distribution, photon distribution analysis (PDA),[67,68] allows us to separate shot noise from inhomogeneous broadening and calculate the FRET-averaged DA distances $\langle R_{DA}\rangle_E$ (for further details on PDA, see section S2.4 in the Supporting Information). A Förster Radius R_0 = 52 Å of Alexa488−Cy5 (assuming an orientation factor $\kappa^2 = 2/3$, justification given in section 3.5) is used to calculate the DA distances with eq 5B. As an example, we present a smFRET histogram of the labeled sample RNA2(19−) in Figure 5B. The PDA analysis clearly shows the presence of the three species: (I) major population (85.9%) of the expected FRET species with $\langle R_{DA}\rangle_E$ = 48.8 Å (orange line); (II) 12.5% of D-only (green line); and (III) 1.6% of impurities with an apparent R_{DA} = 68.8 Å (blue line; also present in donor-only samples). Moreover, a fixed DA distance is not sufficient for the FRET species, and a Gaussian distance distribution with an apparent distribution half width (σ_{app}) has to be used instead. The recovered σ_{app} is about 4−5% of the mean distance $\langle R_{DA}\rangle_E$ and can be attributed mainly to acceptor photophysics.[61] Thus, σ_{app} must not be confused with real physical distance distributions (modeled or recovered by donor fluorescence lifetime analysis of eTCSPC). However, additional broadening leading to considerably larger values for σ_{app} can be due to distance heterogeneities as seen later (section 3.4).

In Figure 5C, the experimental FRET-averaged distances $\langle R_{DA}\rangle_E$ of all five molecules are plotted versus the calculated values for R_{mp} (open symbols) and $\langle R_{DA}\rangle_E$ (full symbols). We assumed a perfect A-RNA (parameters are given in Table S3 in the Supporting Information) for the AV model. As we do not need to find an unknown target structure, it is sufficient to calculate theoretical values of $\langle R_{DA}\rangle_E$ by an $R_{mp}\leftrightarrow\langle R_{DA}\rangle_E$ conversion function (algorithm 1).

We first calculated the coordinates of the mean D- and A-positions, which are given in Table 3 (alternative representations of $\langle \boldsymbol{R}_D\rangle$ and $\langle \boldsymbol{R}_A\rangle$ are given in Table S3B in the Supporting Information). Next, R_{mp} values were calculated and converted to $\langle R_{DA}\rangle_E$ by using the $R_{mp}\leftrightarrow\langle R_{DA}\rangle_E$ conversion function given in the caption shown in Figure S4 (Supporting Information). In Figure 5C, the solid line with a slope of 1 indicates perfect agreement between theory and experiment. Notably, only FRET-averaged (eq 5) DA distances $\langle R_{DA}\rangle_E$ describe the FRET experiment correctly for the whole distance range; that is, the theory must include the distribution of dye positions. Especially for short distances, there are large deviations between the measured FRET distances ($\langle R_{DA}\rangle_E$) and the distances between the modeled mean positions of the dyes (R_{mp}, □), whereas the theoretical and experimental values for $\langle R_{DA}\rangle_E$ agree very well, which is also indicated by an rmsd value of 1.3 Å. The agreement is even better than expected for the given position and orientational uncertainties (dashed lines in Figure 5C; see Table 3 and section 3.5), which suggests that for the five samples studied here κ^2 is very close to 2/3, as discussed in section 3.5. Moreover, the specific linker displacement effect of the dye relative to C5 of the pyrimidine base is described correctly. This effect results in an average z-shift of the mean position of the dye of ∼4.3 Å ($=1/2\Delta z_{link} + 1/2\Delta z_{C5}$) toward the 3′ end (Figure 5C). Thus, simplified modeling of dye positions by the AV method and the described correction for systematic errors is usually sufficient and agrees well with both MD and our experimental data obtained for DNA and RNA.

So far, we have neglected that the refractive index of the macromolecule should be taken into account for Förster radius calculations.[69,70] This small effect may become necessary to consider as the accuracy of FRET increases to angström resolution.

3.3. Influence of DA Distance Distributions on Ensemble Time-Resolved FRET Measurements. In the case of short DA distances with high FRET, uncertainties in dye position represent a large fraction of the absolute R_{DA} values, which may lead to significant systematic errors. As shown in Table 3, the "effective" linkage lengths ($|\langle \boldsymbol{R}_D\rangle|$ and $|\langle \boldsymbol{R}_A\rangle|$) are expected to range typically between 8 and 10 Å. Moreover, distance distribution half widths become comparable to R_{DA}. To characterize the resulting R_{DA} distributions also experimentally, we measured donor and acceptor fluorescence decays in the presence of FRET for labeled DNA and RNA samples with small separations between the dyes (10 and 11 bps for DNA1 and RNA1, respectively), by using eTCSPC. Experimental eTCSPC data in Figure 6 obtained for DNA1 samples could not be fitted using a single FRET rate, that is, with a single fluorescence lifetime to describe the FRET state (see Figure 6, upper residuals plot). We attribute the complex donor decay to a distribution of donor−acceptor distances, which is likely mainly due to the flexibility of the dye linkers. In view of the averaging regime, which has been discussed in section 3.2.2, fluorescence lifetime measurements allow one to obtain snapshots of heterogeneities, which live longer than the fluorescence lifetime (in our case 4 ns). In other words, due to its high time resolution, TCSPC can be used to characterize DA distance distributions $p(R_{DA})$. For example, the analysis of donor fluorescence decays $F_D(t)$ (D decay) by eq 6 recovers $p(R_{DA})$.[7,24,71]

$$F_D(t) = \int_{R_{DA}} p(R_{DA}) \exp\left(-\frac{t}{\tau_{D(0)}}[1+(R_0/R_{DA})^6]\right) dR_{DA} \quad (6)$$

where $\tau_{D(0)}$ is the donor fluorescence lifetime without acceptor. For simplicity, the distribution $p(R_{DA})$ can be assumed to be Gaussian. Considering also the presence of donor-only molecules, the fitting parameters of eq 7 are then the mean DA distance $\langle R_{DA}\rangle$, the half-width σ_{DA} of the R_{DA} distribution, and also the fraction of donor-only molecules x_D (in our measurements typically below 10%):

$$F_D(t) = (1-x_D)\int_{R_{DA}} \frac{1}{\sqrt{2\pi}\sigma_{DA}} \exp\left(-\frac{(R_{DA}-\langle R_{DA}\rangle)^2}{2\sigma_{DA}^2}\right) \exp\left(-\frac{t}{\tau_{D(0)}}[1+(R_0/R_{DA})^6]\right) dR_{DA} + x_D \exp\left(-\frac{t}{\tau_{D(0)}}\right) \quad (7)$$

To test the accuracy of the AV model, we compared the mean and the half-width of the distribution $p(R_{DA})$ obtained by fitting eq 7 to experimental TCSPC data, with those predicted by the AV simulations. A good agreement between experimental and simulated data was found (see section 3.3.1). To justify the use of a Gaussian distribution in eq 7, we simulated a DA distribution of a FRET experiment using two AV position distributions (see Figure S5A in the Supporting Information). Interestingly, this DA distribution $p(R_{DA})$ is very well described by a Gaussian distribution. The validity of the approximation of $p(R_{DA})$ by a single Gaussian distribution can be checked experimentally as demonstrated in section 3.4.

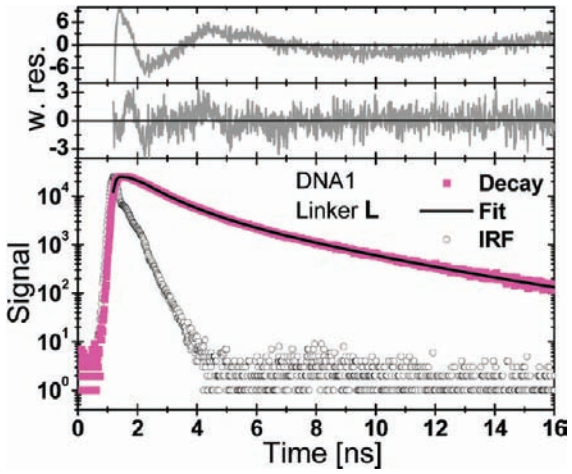


Figure 6. eTCSPC measurement of the fluorescence decay of Alexa488 obtained for the DNA1 FRET sample with L linker. Experimental data (magenta ■), instrument response function (IRF, ○), and the fit assuming a Gaussian distribution of distances (eq 7; bottom panel, solid black line) are shown. Weighted residuals are presented in the upper plots; (top) formal biexponential fluorescence decay with a single FRET state corresponding to $\tau_{D(A)}$ = 1.0 ns (87.5%), and a donor-only decay with $\tau_{D(0)}$ = 4.1 ns (12.5%); χ_r^2 = 10.3; (middle) Gaussian distribution of distances (parameters are given in Table 4) and donor-only decay (eq 7) with $\tau_{D(0)}$ = 4.1 ns (6.7%); χ_r^2 = 1.35. The fit ranges from the maximum of the IRF to the first time channel with less than 100 detected photons.

The extension of eqs 6 and 7 to the case of multiexponential fluorescence relaxation of D-only can be easily made (see ref 72 and section S2.6 in the Supporting Information).

Moreover, the characteristic rise time of the acceptor fluorescence (A rise) also contains information on the FRET rates[72−74] (for details, see section S2.6 and Figure S5B in the Supporting Information).

3.3.1. Measured $\langle R_{DA}\rangle$ Distances Do Not Compare to Simulated R_{mp}. The simulated and measured distances are presented in Table 4. Clearly, the values for distances between mean dye positions (R_{mp}) simulated by the AV method do not agree with the distances measured by eTCSPC and PDA. To explain these deviations, one should note that, like for $\langle R_{DA}\rangle_E$ (section 3.2.2), the mean distance $\langle R_{DA}\rangle$ measured by eTCSPC (eqs 6 and 7) is not equal to R_{mp} but is rather given by eq 8:

$$\langle R_{DA}\rangle = \langle |\boldsymbol{R}_D - \boldsymbol{R}_A| \rangle \geq |\langle \boldsymbol{R}_D\rangle - \langle \boldsymbol{R}_A\rangle| \qquad (8)$$

In other words, $\langle R_{DA}\rangle \geq R_{mp}$, so that a difference of >5 Å can be seen for broader distance distributions (Table 4). For instance, the distance between the centers of completely overlapping dye clouds is zero (R_{mp} = 0), whereas the average distance between individual dye positions is not ($\langle R_{DA}\rangle$ > 0). Thus, a realistic modeling of dye position distributions is needed not only for PDA but also for a proper interpretation of time-resolved fluorescence decays. If one compares $\langle R_{DA}\rangle$ with $\langle R_{DA}\rangle_E$, the latter is typically weighted toward R_0, that is, $\langle R_{DA}\rangle_E \geq \langle R_{DA}\rangle$ for $\langle R_{DA}\rangle < R_0$ and $\langle R_{DA}\rangle_E \leq \langle R_{DA}\rangle$ for $\langle R_{DA}\rangle > R_0$.

Analysis of DNA. For DNA samples with L and I linker, the agreement between the simulated and experimental $\langle R_{DA}\rangle$ values obtained by eTCSPC and PDA is good, and remaining deviations are well within the expected position (Table 3) and the κ^2-related errors (section 3.5). $\langle R_{DA}\rangle$ and σ_{DA} were also extracted by fitting the acceptor rises (Table 4). The good agreement between

Table 4. DA Distance Distribution Parameters Calculated from eTCSPC and smFRET Data in Comparison with AV Simulations[a]

	eTCSPC experiment		PDA	AV simulation	
linker	$\langle R_{DA}\rangle$ (σ_{DA}), Å (D decay)	$\langle R_{DA}\rangle$ (σ_{DA}), Å (A rise)	$\langle R_{DA}\rangle_E$ (σ_{app}), Å	R_{mp}, Å	$\langle R_{DA}\rangle$ (σ_{DA}), Å
DNA1					
L	39.5 (7.5)	42.3 (6.8)	42.5 (1.5)	35.5	39.9 (8.8)
I	40.8 (5.9)	40.8 (8.2)	41.5 (1.5)	33.9	36.3 (6.0)
S_t	38.4 (11.2)	45.6 (7.7)	45.9 (2.9)	32.7	35.6 (4.8)
RNA1					
L	41.5 (7.3)	43.2 (6.3)	42.8 (1.6)	27.5	31.6 (6.3)
S_d	35.1 (6.0)	34.5 (6.8)	36.6 (1.0)	22.7	23.9 (5.1)
S_t	34.4 (7.6)	34.6 (11.2)	36.5 (4.7)	23.0	24.4 (5.3)

[a] The dye−dye separation is 10 and 11 basepairs for DNA and RNA, respectively.

the values calculated from donor and acceptor decay curves shows that donor quenching is indeed due to FRET and not due to, for example, local quenching artifacts. For S_t linkers, systematic deviations are seen, and the R_{DA} distributions are surprisingly broad (σ_{DA} = 11.2 Å), which is clearly unexpected for the shortest linker. We will show below that a rather irregular distribution of stable conformations of one or both dyes must exist for S_t linkers, which cannot be described with a single Gaussian peak as assumed in eq 7.

Analysis of RNA. In the case of RNA, a similarly broad R_{DA} distribution can be observed for the acceptor rise for the S_t linker. Furthermore, the agreement between theory and experiment is generally worse than for DNA. The main reason is that the DA distances are quite short (less than one-half of the Förster radius of 52 Å). At such short distances, the point dipole approximation does not hold[8,20,75,76] and should be replaced accordingly, which is beyond the scope of the work. Furthermore, dye−dye interactions[77] cannot be excluded. However, let us point out that the measurements with short linkers (S_d or S_t) provide distances that are much closer to the structurally relevant distance between the C5 atoms of the labeled nucleobases (30 Å) than for L linkers.

3.3.2. Translational Linker Movements Are Slow on the Time Scale of FRET. eTCSPC data provide additional evidence for quasi-static distance distributions on the time scale of FRET as has been postulated before.[25] Considering L, I, and S_d linkers, the fit of R_{DA} distributions to experimental fluorescence decays typically yields distribution half-widths σ_{DA} comparable to simulated ones (Table 4). This indicates that DA distributions due to linker motions are not significantly averaged out on the nanosecond time scale. There is no contradiction between this finding and subns anisotropy decay times: for example, Alexa488 readily rotates about the linker axis, which involves strong fluorescence depolarization but very little distance fluctuations. In the case of low FRET, there might be more averaging of distances; however, taking into account large absolute values of R_{DA}, only minor relative errors could be expected because the differences in E for dynamic and static averaging are small for large distances (for low FRET, $\langle R_{DA}\rangle \approx R_{mp}$ irrespective of the averaging regime).

3.4. FRET Broadening on the Millisecond Time Scale. To understand the surprisingly broad R_{DA} distributions for the S_t

linkers (section 3.3), we performed single-molecule experiments on freely diffusing DNA1 and RNA1 molecules to check whether the distributions are averaged out on the millisecond time scale. Figure 7A−D shows 2D probability histograms of FRET efficiency E versus the donor lifetime $\tau_{D(A)}$, where the corresponding 1D parameter histograms are given as projections. The number of molecules (fluorescence bursts) in each bin is gray scale shaded from white (lowest) to black (highest). All 2D plots show two distinct peaks: (I) at $E \approx 0$ and $\tau_{D(A)} \approx 4$ ns, due to D-only populations, and (II) at $E \approx 0.7-0.8$ and $\tau_{D(A)} \approx 1-2$ ns attributed to FRET subpopulations. In all plots, solid lines indicate the expected E versus $\tau_{D(A)}$ dependence for dynamic FRET taking fast linker dynamics into account[78] (eqs S12 and S13 in section S2.6 of the Supporting Information). DNA and RNA samples with L linkers show homogeneous uncorrelated $E-\tau_{D(A)}$ distributions of the FRET population (Figure 7A and B), which are distributed approximately horizontally due to independent shot noise distributions of E[68] and $\tau_{D(A)}$.[79] All samples with I and S_d linkers exhibit similar patterns (see Figure S6 for RNA1 with S_d linker in section S2.7 of the Supporting Information).

On the contrary, DNA1 and RNA1 with S_t linkers have broad and asymmetric 2D distributions of $E-\tau_{D(A)}$ (Figure 7C and D). Moreover, for S_t linkers, a correlation between E and $\tau_{D(A)}$ within the FRET subpopulation is apparent from 2D plots, indicating a heterogeneous distribution of DA distances on the millisecond time scale. This fact is consistent with the eTCSPC data (Table 4), which show unexpectedly broad R_{DA} distributions obtained exclusively for S_t linkers. The comparison of two complementary ds RNA1 FRET samples with different linkers for D and A, respectively, yields that this is mainly caused by the donor. Only the sample labeled with linker S_t at the donor and linker L at the acceptor position shows E broadening, whereas E is narrow for the RNA1 sample with the opposite linker combination of the dyes (Figure S7 in section S2.7 of the Supporting Information).

PDA. To quantify these visual effects, we performed PDA of the smFRET data shown in Figure 7A−D. The mean FRET-averaged DA distances $\langle R_{DA}\rangle_E$ recovered by PDA are typically similar to $\langle R_{DA}\rangle$ values found by eTCSPC (Table 4). A large deviation is observed only for DNA1 with S_t linkers, indicating distinct dye subpopulations. Moreover, the apparent R_{DA} distribution half-widths found by PDA (σ_{app}) are significantly larger for S_t linkers (Table 4). For instance, for RNA1, σ_{app} amounts ~13% of $\langle R_{DA}\rangle_E$ for S_t linkers, whereas for other linkers σ_{app} does not exceed 3−4% of $\langle R_{DA}\rangle_E$. This fact indicates the presence of FRET heterogeneities in addition to the complex acceptor's photophysics, which is responsible for the minimal broadening of few percent of $\langle R_{DA}\rangle_E$.[61]

Maximum Entropy Deconvolution. To further support this conclusion, we reanalyzed all data with S_t linkers and used an unbiased model-free DA distance distribution instead of the assumed Gaussian distribution. We performed a maximum entropy (ME) deconvolution[47,48] of R_{DA} distributions observed for S_t linkers on different time scales (Figure 7E and F): eTCSPC data for the nanosecond time scale (dashed lines) and PDA of smFRET intensity data for the millisecond time scale (solid lines). ME deconvolutions of both experiment types indicate the presence of multiple FRET states. eTCSPC data show that S_t linkers have two stable conformations (corresponding to two major "stable" peaks "s1" and "s2" in Figure 7E and F) possibly with some additional flexibility leading to broadening of these peaks. For RNA1, PDA data show a peak between the two DA distances found by eTCSPC ("mixing" peak "m" in Figure 7F), which suggests that additional averaging takes place on a time scale faster than the dwell time (ms). Dynamic PDA[78,80] provides no indication for "mixing" with millisecond characteristic times (see Figure S8 in the Supporting Information). Thus, transitions responsible for the "mixing" peak occur on a significantly faster time scale, whereas the "s1" and "s2" peaks are stable during the burst duration.

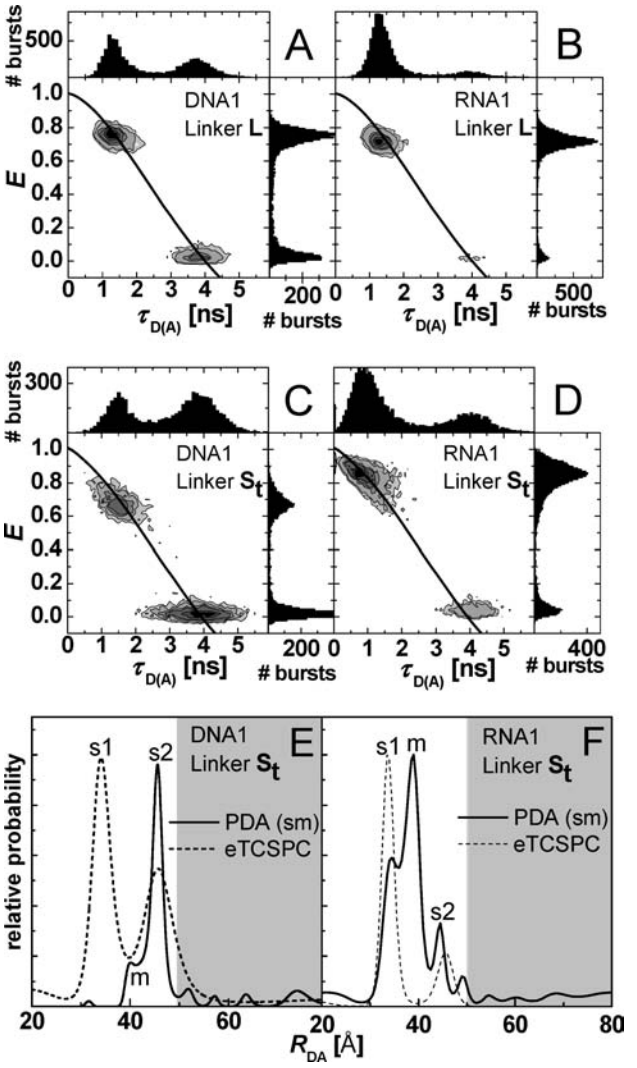


Figure 7. (A−D) 2D probability histograms of FRET efficiency E versus the donor lifetime $\tau_{D(A)}$ generated from smFRET data obtained for (A) DNA1, L linkers; (B) RNA1, L linkers; (C) DNA1, S_t linkers; and (D) RNA1, S_t linkers. The number of molecules (fluorescence bursts) in each bin is gray scale shaded from white (lowest) to black (highest). The corresponding 1D parameter histograms are given as projections. In all plots, solid lines indicate the E versus $\tau_{D(A)}$ dependence given by eqs S12 and S13 in section S2.6 of the Supporting Information. (E,F) Maximum entropy deconvolution of donor−acceptor distance distributions for (E) DNA1, S_t linkers and (F) RNA1, S_t linkers. The solid and dashed lines represent deconvolutions of single-molecule intensity data and donor decays measured by eTCSPC, respectively. The ME method allows one to extract distributions of fluorescence lifetimes from eTCSPC data,[48] which can be converted into distributions of R_{DA}. In PDA,[68] distance ME distributions are directly used to calculate FRET efficiency distributions. The shaded areas indicate the region where the observed peaks may represent multimolecular and photobleaching events. The peaks "s1" and "s2" correspond to stable dye conformations, whereas the peak "m" is likely due to averaging of the states s1 and s2 on the millisecond time scale.

For DNA1, a significant discrepancy between PDA and eTCSPC data is observed (cf., Figure 7E and Tables 1 and 4). This effect could be due to a static quenching of the acceptor in the state corresponding to the shorter R_{DA} (peak "s1" in Figure 7E). If so, the s1 state would be nearly invisible in SMD measurements because both dyes are quenched in this state (D is quenched by high FRET), which also explains the disagreement between donor decay and acceptor rise data (Table 4). This explanation is supported by ensemble measurements of the acceptor fluorescence quantum yield Φ_{FA}, which is reduced for the S_t linker ($\Phi_{FA} = 0.21$) as compared to the L linker ($\Phi_{FA} = 0.38$, Table 1).

In summary, complex distributions of DA distances found for S_t linkers severely complicate their use in quantitative FRET studies. Thus, an optimal linker must always allow for free diffusion within the dye's accessible volume. On the other hand, to minimize position uncertainties and to prevent dye−dye interactions, linkers should not be longer than necessary. For DNA and Alexa488, the I linker seems to fit best to these requirements. For Cy5, shorter S_d and S_t linkers work equally well because the dye has an extra 6-atom linker between the chromophore and the reactive coupling group (Figure 2). Hence, for undefined dye environments, where modeling of dye positions is impossible, I and S_d are well suitable linkers as they provide better absolute distance estimations.

3.5. Minimizing Uncertainties Due to the Orientational Factor κ^2. For accurate FRET analysis, we must consider not only translational linker diffusion, which is usually slower than ns, but also orientational dynamics. Given subnanosecond local rotational correlation times in Table 2, dynamic orientational averaging can be assumed at least to some extent. Dye reorientation dynamics on the time scale of FRET (k_{FT}, see section 3.2.2)[81] could, thus, be relevant only for very short distances, which is beyond the scope of this Article. The anisotropy measurements indicate, however, that the orientational distribution of both D and A is not strictly isotropic even for L-linkers (Table 2), contrary to the MD data from[25] (for details, see section S2.8 in the Supporting Information). Therefore, detailed analysis of related κ^2 effects for all linkers must be performed.

As mentioned above, the modeling of the accessible space of a dye is much less important for short linkers, making them more suitable for quantitative FRET measurements in the case of undefined environments. One potential problem when using short linkers, however, is the estimation of the orientation factor κ^2. The assumption of $\kappa^2 \cong 2/3$ might not be justified for short linkers as their movement is more restricted. In this section, we estimate the range of possible values for κ^2 [19,82,83] for the different linkers of DNA1 and RNA1 and determine potential errors for FRET distance measurements if $\kappa^2 = 2/3$ is assumed in the calculation of R_0.

To minimize the uncertainty in κ^2, we use the residual anisotropies that result from the measurements of the donor-only, the acceptor-only (Figure 3B in Section 3.1.2.), and the FRET-sensitized acceptor (Figure 8A) anisotropy decays ($r_{\infty,D}$, $r_{\infty,A}$, and $r_{\infty,A(D)}$, respectively), to calculate the range of all possible values for κ^2 by eq 9:

$$\kappa^2 = \frac{2}{3} + \frac{2}{3}S_D^{(2)}S^{(2)}(\beta_1) + \frac{2}{3}S_A^{(2)}S^{(2)}(\beta_2) + \frac{2}{3}S_D^{(2)}S_A^{(2)}(S^{(2)}(\delta) + 6S^{(2)}(\beta_1)S^{(2)}(\beta_2) + 1 + 2S^{(2)}(\beta_1) + 2S^{(2)}(\beta_2) - 9\cos\beta_1\cos\beta_2\cos\delta) \quad (9)$$

In eq 9, β_1 and β_2 are the angles between the symmetry axes of the rotations of the dyes and the distance vector $\boldsymbol{R}_{DA}$, while δ is the angle between the symmetry axes (Figure 8B).

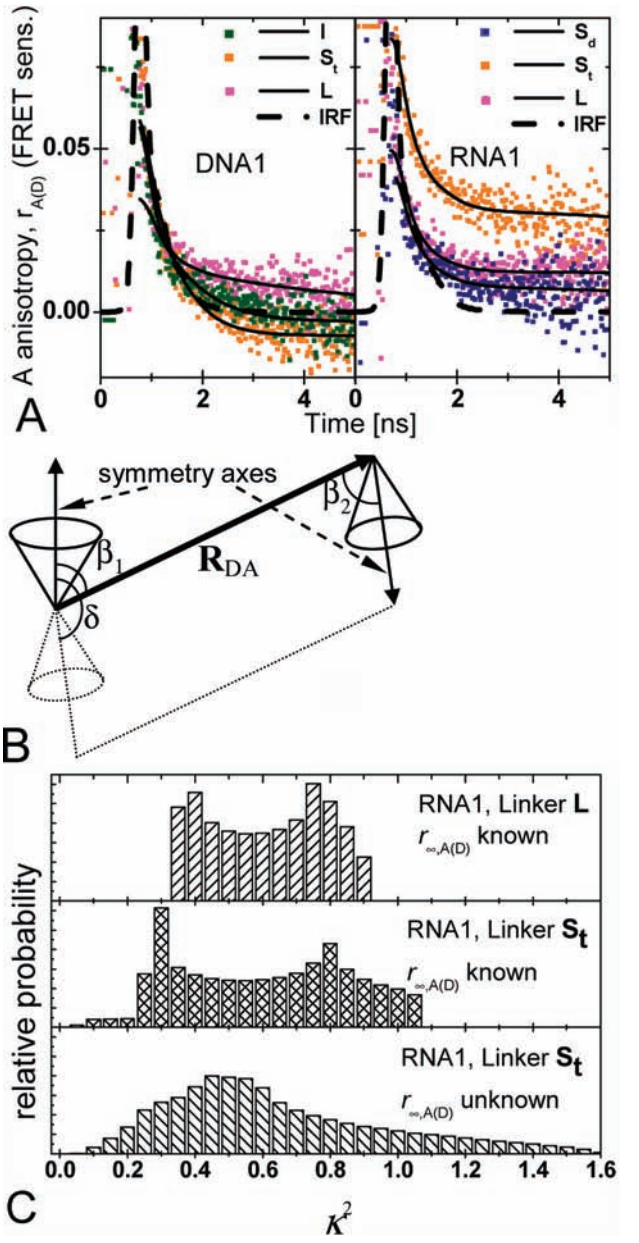


Figure 8. (A) FRET-sensitized acceptor anisotropy decays (dots), the rescaled instrument response functions (IRF, dashed line), and the anisotropy decays (black solid lines) of DNA (left) and RNA (right) (fit results, see Table 5). (B) Sketch showing angles that define the orientational factor κ^2 in eq 9: β_1 and β_2 are the angles between the symmetry axes of the rotations of the dyes and the distance vector $\boldsymbol{R}_{DA}$, and δ is the angle between the symmetry axes. (C) Calculated probability distributions for possible values of κ^2 in RNA1: (top) for linker L, (middle) for linker S_t and in both cases taking into account $r_{\infty,A(D)}$; (bottom) for linker S_t and only using the offsets from the anisotropy decays of the donor and acceptor molecules $r_{\infty,D}$ and $r_{\infty,A}$.

The necessary second-rank order parameters $S^{(2)}$ are defined in eq 3 ($S_D^{(2)}$) and eq 4 ($S_A^{(2)}$) and formally by:

$$S^{(2)}(\delta) = \frac{1}{2}(3\cos^2\delta - 1) = \frac{r_{\infty,A(D)}}{r_0 S_D^{(2)} S_A^{(2)}},\ S^{(2)}(\beta_1) = \frac{1}{2}(3\cos^2\beta_1 - 1)\ ,\ \text{and}\ S^{(2)}(\beta_2) = \frac{1}{2}(3\cos^2\beta_2 - 1) \quad (10)$$

dx.doi.org/10.1021/ja105725e | J. Am. Chem. Soc. 2011, 133, 2463–2480

Table 5. Fit Parameters for FRET-Sensitized Acceptor Anisotropy Decay ($r_0 = 0.38$) and Results for κ^2 Probability Distributions Taking into Account $r_{\infty,D}$, $r_{\infty,A}$ (Table 2), and $r_{\infty,A(D)}$

	sensitized anisotropy			border values: worst cases			mean value: typical case		
linker	r_S	ρ_1, ns (b_1)	$\rho_{2(global)}$, ns ($b_2 = r_{\infty,DA}$)	κ^2_{min}	κ^2_{max}	ΔR_{DA}	κ^2_{mean}	accuracy of R_{DA}	precision ΔR_{DA}
				DNA1					
I	0.005	0.605 (0.042)	>60 (−0.003)	0.25	0.94	−6%...+15%	0.6	+3.1%	7.0%
S_t	0.002	0.400 (0.080)	>60 (−0.007)	0.22	0.92	−6%...+17%	0.59	+2.6%	5.9%
L	0.012	0.227 (0.056)	10 (0.010)	0.42	0.92	−6%...+ 7%	0.67	+1.7%	4.4%
				RNA1					
S_d	0.012	0.260 (0.045)	>60 (0.007)	0.10	1.57	−13%...+37%	0.64	+0.8%	8.4%
S_t	0.038	0.258 (0.069)	>60 (0.031)	0.09	1.1	−9%...+28%	0.57	+3.9%	8.9%
L	0.016	0.246 (0.047)	>60 (0.012)	0.38	0.93	−6%...+ 9%	0.62	+2.2%	5.1%

κ^2 values (eq 9) cannot be calculated unambiguously because, in general, the angles β_1 and β_2 are not experimentally accessible. However, for an experimentally determined δ (eq 10), it is possible to define a range of possible values for β_1 and β_2 (eqs S16A and S16B in section S2.8 of the Supporting Information), which allows one to calculate the range of κ^2 values (eq 9) compatible with the experimental data ($r_{\infty,D}$, $r_{\infty,A}$, and $r_{\infty,A(D)}$).

Note that κ^2 does not explicitly depend on the cone opening half angles θ_D and θ_A (section 3.1.2) and the assumption of dye reorientation within a cone/disk (Figure 3C and D); that is, even if this approximation is considered as unrealistic, it is not needed to obtain eq 9. However, axially symmetric transition dipole orientation distributions are usually assumed for κ^2 estimation,[19] which might be not exactly the case here (Figure 4). Therefore, it is absolutely reasonable to expect that additional averaging of mutual orientations of D and A by diffusion along the DNA or RNA groove would bring the effective κ^2 even closer to 2/3 than given by eq 9. This might be one of the reasons for the (unexpected) high accuracy of our FRET benchmark study using RNA2 and RNA3 (Figure 5C in section 3.2.2). Thus, the κ^2 estimations given below represent the worst case scenario. Moreover, in this study, we cannot investigate a possible correlation between R_{DA} and κ^2 as proposed in a few recent theoretical works[20,75,76] for the case that the dye exhibits a slow exchange between different microenvironments. Our simplified (AV) simulations do not allow us to discuss this effect.

3.5.1. Estimation of κ^2 Using $r_{\infty,A(D)}$. The residual anisotropies $r_{\infty,D}$, $r_{\infty,A}$, and $r_{\infty,A(D)}$ can be used to calculate a probability distribution of κ^2 values. For the estimation of $r_{\infty,A(D)}$, the FRET-sensitized acceptor anisotropy decays were studied for all linkers of DNA1 and RNA1 (see section 2 for the measurement and analysis procedures). The data and the results of the analysis are shown in Figure 8A and Table 5, respectively. In all cases, the decays could be fitted by a biexponential decay consisting of a fast decay time $\rho_1 \approx 0.4$ ns (resulting from fast FRET (k_{FT}) and local reorientations (k_R)) and a slow decay time ρ_2 (global motion). As in section 3.1, the slow component is approximated by a time-independent offset $r_{\infty,A(D)}$ and considered to be the residual anisotropy. It is largest for the linker S_t in RNA and is negative for the linkers I and S_t in DNA, which indicates that the transition dipole moments of the dyes are preferentially orientated perpendicular to each other. Indeed, the differences between all anisotropy decays indicate distinct mean dye orientations. Moreover, smFRET measurements were analyzed by inspecting 2D probability histograms of steady state anisotropy r_s versus the donor lifetime $\tau_{D(A)}$ to make sure that the resulting residual anisotropies $r_{\infty,A(D)}$ are due to restricted reorientation of the dyes and not due to a fraction of molecules where the dye is immobile. For linkers L, the FRET populations in the $r_s-\tau_{D(A)}$ 2D plots appear to be symmetric at $r_s \approx 0.1$ and $\tau_{D(A)} \approx 1.3$ ns. Thus, there is only one anisotropy population for each lifetime population. For shorter linkers, the anisotropy distributions become less symmetric. The existence of additional donor populations is most pronounced for S_t linkers with a fraction of a 2-fold increased steady-state anisotropy, which amounts to less than 29% (see Figure S9 in section S2.8 of the Supporting Information for all FRET pairs). However, in no case could a completely immobile dye species be detected.

Three distributions of possible κ^2 values are presented in Figure 8C with the corresponding parameters compiled in Table 5. The smallest and the largest possible values κ^2_{min} and κ^2_{max} represent the worst case scenario for a FRET distance measurement for which $\kappa^2 = 2/3$ is assumed. The resulting range of potential relative errors (worst case) for distance measurements ΔR_{DA} typically vary from −6% to +15% (Table 5). For a range of possible κ^2 values, one can define an accuracy (systematic error) and a precision (uncertainty; eqs S17A and S17B in section S2.8 of the Supporting Information), which would result in a typical uncertainty of κ^2 in the FRET distance measurements presented here. These parameters are shown in Table 5, and it is obvious that the errors and deviations are relatively small in all cases, although they become larger for short linkers (see also Figure 8C). Nevertheless, the distances measured using short linkers are reasonably accurate.

3.5.2. Estimation of κ^2 with Unknown $r_{\infty,A(D)}$. For the estimation of κ^2, many groups only use $r_{\infty,D}$ and $r_{\infty,A}$,[7] which result from the anisotropy decays of the D-only and the A-only molecules, respectively, and do not measure the FRET sensitized acceptor anisotropy decay. Hence, $r_{\infty,A(D)}$ is unknown, and δ must be allowed to be every value between 0° and 90°. For comparison, the probability distributions for κ^2 for all linkers of DNA1 and RNA1 were also calculated accordingly; that is, only $r_{\infty,D}$ and $r_{\infty,A}$ were taken into account. The results are shown in Table S4 in section S2.8 of the Supporting Information. There are only slight changes in accuracy and precision, but the range of possible κ^2 values nearly doubles. Thus, the worst-case ΔR_{DA} increases especially toward smaller distance values from typically ∼ −6% for known $r_{\infty,A(D)}$ (section 3.5.1) to typically −15% for unknown $r_{\infty,A(D)}$. Figure 8C shows the probability distributions for the RNA molecules for the linkers L and S_t in case all three residual

anisotropies are taken into account (top and middle, respectively) and for linker S_t when only $r_{\infty,D}$ and $r_{\infty,A}$ are used (bottom).

It is worth mentioning that for the estimation of κ^2 the steady-state anisotropy r_s is commonly used instead of the residual anisotropy r_∞.[7,29] This method, however, overestimates κ^2-related errors even further as r_s is usually larger than r_∞.

4. CONCLUSION

In this work, we introduce new short dye linkers for labeling of DNA and RNA. FRET measurements on dsDNA and dsRNA model systems test their suitability for quantitative studies. For well-defined environments and if the dye diffuses freely in the sterically allowed, mean positions of the dyes and R_{DA} distributions can be accurately modeled using relatively simple and fast accessible volume simulations. The translational motions of the dyes appear to be slow on the time scale of the fluorescence lifetimes and depend on linker size and structure. For L, I, and S_d linkers, the motions are completely averaged out on the millisecond time scale, whereas S_t linkers exhibit a complex distribution of R_{DA} due to inhibited diffusion through the accessible volume. It became also clear that experimentally measured distances, especially for long linkers, cannot be directly compared to the structurally relevant mean position distance R_{mp} or the distance between C5 atoms of uracil. Because of broad distributions of DA distances, there is a large discrepancy between R_{mp} and $\langle R_{DA}\rangle_E$ (measured by smFRET) or $\langle R_{DA}\rangle$ (measured by eTCSPC), which must be always taken into account. When doing so, high precision distance measurements are possible. This correction and, therefore, modeling of dye positions is less important for short linkers. This makes them particularly suitable for undefined environments. However, due to the inhomogeneities of DA distances observable for S_t as a linker for Alexa488, we advise against its use and recommend the S_d or the I linker instead.

It became clear that, when calculating a probability distribution for possible values of κ^2, its width can be further reduced when taking into account not only the residual anisotropies of the donor and acceptor $r_{\infty,D}$ and $r_{\infty,A}$ but also $r_{\infty,A(D)}$, which results from the FRET-sensitized acceptor anisotropy decay. Furthermore, for short and intermediate linkers, κ^2-related errors are only slightly higher than for long linkers. When using them for quantitative FRET measurements of internally labeled nucleic acids with Alexa488 and Cy5 as a FRET pair, it is, therefore, safe to assume κ^2 to be 2/3. For unknown local environments, long linkers increase uncertainties significantly more.

■ ASSOCIATED CONTENT

Supporting Information. Additional experimental procedures, DNA and RNA sample information, and additional results. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

stanislav.kalinin@uni-duesseldorf.de; sabine.mueller@uni-greifswald.de; cseidel@hhu.de

Author Contributions

‖These authors contributed equally.

■ ACKNOWLEDGMENT

C.A.M.S., S.M., H.N., and S.K. thank the German Science foundation (DFG) in the priority program SPP 1258 "Sensory and regulatory RNAs in prokaryotes" for funding this work. S.S. thanks the NRW Research School Biostruct for funding. We are grateful to Hayk Vardanyan for help in single-molecule measurements. We thank Evangelos Sisamakis, Hugo Sanabria, Ralf Kühnemuth, and Suren Felekyan for very fruitful discussions. We thank Anna K. Woźniak for providing the cartoon for the TOC graphic.

■ REFERENCES

(1) Stryer, L.; Haugland, R. P. *Proc. Natl. Acad. Sci. U.S.A.* **1967**, *58*, 719–726.
(2) Clegg, R. M. *Methods Enzymol.* **1992**, *211*, 353–388.
(3) van der Meer, B. W.; Cooker, G.; Chen, S. Y. *Resonance Energy Transfer: Theory and Data*; VCH Publishers: New York, 1994.
(4) Sisamakis, E.; Valeri, A.; Kalinin, S.; Rothwell, P. J.; Seidel, C. A. M. *Methods Enzymol.* **2010**, *475*, 456–514.
(5) Roy, R.; Hohng, S.; Ha, T. *Nat. Methods* **2008**, *5*, 507–516.
(6) Förster, T. *Ann. Phys.* **1948**, *437*, 55–75.
(7) Lakowicz, J. R. *Principles of Fluorescence Spectroscopy*; Kluwer Academic/Plenum Publishers: New York, 1999.
(8) Valeur, B. *Molecular Fluorescence: Principles and Applications*; Wiley-VCH Verlag: Weinheim, 2002.
(9) Scholes, G. D. *Annu. Rev. Phys. Chem.* **2003**, *54*, 57–87.
(10) Ha, T.; Enderle, T.; Ogletree, D. F.; Chemla, D. S.; Selvin, P. R.; Weiss, S. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 6264–6268.
(11) Deniz, A. A.; Dahan, M.; Grunwell, J. R.; Ha, T. J.; Faulhaber, A. E.; Chemla, D. S.; Weiss, S.; Schultz, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 3670–3675.
(12) Rothwell, P. J.; Berger, S.; Kensch, O.; Felekyan, S.; Antonik, M.; Wöhrl, B. M.; Restle, T.; Goody, R. S.; Seidel, C. A. M. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 1655–1660.
(13) Margittai, M.; Widengren, J.; Schweinberger, E.; Schröder, G. F.; Felekyan, S.; Haustein, E.; König, M.; Fasshauer, D.; Grubmüller, H.; Jahn, R.; Seidel, C. A. M. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 15516–15521.
(14) Sanborn, M. E.; Connolly, B. K.; Gurunathan, K.; Levitus, M. *J. Phys. Chem. B* **2007**, *111*, 11064–11074.
(15) Neubauer, H.; Gaiko, N.; Berger, S.; Schaffer, J.; Eggeling, C.; Tuma, J.; Verdier, L.; Seidel, C. A. M.; Griesinger, C.; Volkmer, A. *J. Am. Chem. Soc.* **2007**, *129*, 12746–12755.
(16) Norman, D. G.; Grainger, R. J.; Uhrin, D.; Lilley, D. M. J. *Biochemistry* **2000**, *39*, 6317–6324.
(17) Vámosi, G.; Gohlke, C.; Clegg, R. M. *Biophys. J.* **1996**, *71*, 972–994.
(18) Ranjit, S.; Gurunathan, K.; Levitus, M. *J. Phys. Chem. B* **2009**, *113*, 7861–7866.
(19) Dale, R. E.; Eisinger, J.; Blumberg, W. E. *Biophys. J.* **1979**, *26*, 161–193.
(20) VanBeek, D. B.; Zwier, M. C.; Shorb, J. M.; Krueger, B. P. *Biophys. J.* **2007**, *92*, 4168–4178.
(21) Iqbal, A.; Arslan, S.; Okumus, B.; Wilson, T. J.; Giraud, G.; Norman, D. G.; Ha, T.; Lilley, D. M. J. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 11176–11181.
(22) Clegg, R. M.; Murchie, A. I. H.; Zechel, A.; Carlberg, C.; Diekmann, S.; Lilley, D. M. J. *Biochemistry* **1992**, *31*, 4846–4856.
(23) Clegg, R. M.; Murchie, A. I. H.; Zechel, A.; Lilley, D. M. J. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 2994–2998.
(24) Haas, E. *ChemPhysChem* **2005**, *6*, 858–870.
(25) Woźniak, A. K.; Schröder, G.; Grubmüller, H.; Seidel, C. A. M.; Oesterhelt, F. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 18337–18342.
(26) Best, R. B.; Merchant, K. A.; Gopich, I. V.; Schuler, B.; Bax, A.; Eaton, W. A. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 18964–18969.
(27) Merchant, K. A.; Best, R. B.; Louis, J. M.; Gopich, I. V.; Eaton, W. A. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 1528–1533.

(28) Cai, Q.; Kusnetzow, A. K.; Hideg, K.; Price, E. A.; Haworth, I. S.; Qin, P. Z. *Biophys. J.* **2007**, *93*, 2110–2117.
(29) Muschielok, A.; Andrecka, J.; Jawhari, A.; Bruckner, F.; Cramer, P.; Michaelis, J. *Nat. Methods* **2008**, *5*, 965–971.
(30) Xie, Y.; Maxson, T.; Tor, Y. *J. Am. Chem. Soc.* **2010**, *132*, 11896–11897.
(31) Borjesson, K.; Preus, S.; El-Sagheer, A. H.; Brown, T.; Albinsson, B.; Wilhelmsson, L. M. *J. Am. Chem. Soc.* **2009**, *131*, 4288–4293.
(32) Lewis, F. D. *Pure Appl. Chem.* **2006**, *78*, 2287–2295.
(33) Agrofoglio, L. A.; Gillaizeau, I.; Saito, Y. *Chem. Rev.* **2003**, *103*, 1875–1916.
(34) Heck, R. F. *Org. React.* **1982**, *27*, 345–390.
(35) Hobbs, F. W., Jr. *J. Org. Chem.* **1989**, 54.
(36) Dey, S.; Sheppard, T. L. *Org. Lett.* **2001**, *3*, 3983–3986.
(37) Muschielok, A.; Michaelis, J. *FRETnps Tools, Version 2.6.1*; Ludwig Maximilians University: Munich, 2008.
(38) Widengren, J.; Kudryavtsev, V.; Antonik, M.; Berger, S.; Gerken, M.; Seidel, C. A. M. *Anal. Chem.* **2006**, *78*, 2039–2050.
(39) Cruickshank, K. A.; Stockwell, D. L. *Tetrahedron Lett.* **1988**, *29*, 5221–5224.
(40) McKeen, C. M.; Brown, L. J.; Nicol, J. T. G.; Mellor, J. M.; Brown, T. *Org. Biomol. Chem.* **2003**, *1*, 2267–2275.
(41) Schaller, H.; Weimann, G.; Khorana, H. G.; Lerch, B. *J. Am. Chem. Soc.* **1963**, *85*, 3821.
(42) Shah, K.; Wu, H. Y.; Rana, T. M. *Bioconjugate Chem.* **1994**, *5*, 508–512.
(43) Hakimelahi, G. H.; Proba, Z. A.; Ogilvie, K. K. *Can. J. Chem.* **1982**, *60*, 1106–1113.
(44) Bannwarth, W. *Chimia* **1987**, *41*, 302–317.
(45) Schmidt, C.; Welz, R.; Muller, S. *Nucleic Acids Res.* **2000**, *28*, 886–894.
(46) O'Connor, D. V.; Phillips, D. *Time-correlated Single Photon Counting*; Academic Press: New York, 1984.
(47) Livesey, A. K.; Skilling, J. *Acta Crystallogr., Sect. A* **1985**, *41*, 113–122.
(48) Brochon, J. C. *Methods Enzymol.* **1994**, *240*, 262–311.
(49) Kühnemuth, R.; Seidel, C. A. M. *Single Mol.* **2001**, *2*, 251–254.
(50) Eggeling, C.; Berger, S.; Brand, L.; Fries, J. R.; Schaffer, J.; Volkmer, A.; Seidel, C. A. M. *J. Biotechnol.* **2001**, *86*, 163–180.
(51) Arden-Jacob, J. Universität-Gesamthochschule, Siegen, 1992.
(52) Seidel, C. A. M.; Schulz, A.; Sauer, M. H. M. *J. Phys. Chem.* **1996**, *100*, 5541–5553.
(53) Eggeling, C.; Fries, J. R.; Brand, L.; Günther, R.; Seidel, C. A. M. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 1556–1561.
(54) Doose, S.; Neuweiler, H.; Sauer, M. *ChemPhysChem* **2009**, *10*, 1389–1398.
(55) Levitus, M.; Ranjit, S. *Q. Rev. Biophys.* **2010**, *44*, 123–151.
(56) Chibisov, A. K.; Zakharova, G. V.; Gorner, H.; Sogulyaev, Y. A.; Mushkalo, I. L.; Tolmachev, A. I. *J. Phys. Chem.* **1995**, *99*, 886–893.
(57) Widengren, J.; Schwille, P. *J. Phys. Chem. A* **2000**, *104*, 6416–6428.
(58) Widengren, J.; Schweinberger, E.; Berger, S.; Seidel, C. A. M. *J. Phys. Chem. A* **2001**, *105*, 6851–6866.
(59) Sauerwein, B.; Murphy, S.; Schuster, G. B. *J. Am. Chem. Soc.* **1992**, *114*, 7920–7922.
(60) Buschmann, V.; Weston, K. D.; Sauer, M. *Bioconjugate Chem.* **2003**, *14*, 195–204.
(61) Kalinin, S.; Sisamakis, E.; Magennis, S. W.; Felekyan, S.; Seidel, C. A. M. *J. Phys. Chem. B* **2010**, *114*, 6197–6206.
(62) Garcia de la Torre, J.; Huertas, M. L.; Carrasco, B. *Biophys. J.* **2000**, *78*, 719–730.
(63) McDowell, S. E.; Spackova, N.; Sponer, J.; Walter, N. G. *Biopolymers* **2007**, *85*, 169–184.
(64) Case, D. A.; et al. *AmberTools, Version 1.3*; University of California: San Francisco, CA, 2008.
(65) DeLano, W. L. *The PyMOL Molecular Graphics System, Version 1.2r1*; DeLano Scientific: San Carlos, CA, 2002.
(66) *Hyperchem (TM) Professional, Version 7.51*; Hypercube, Inc.: Gainesville, FL, 2003.
(67) Antonik, M.; Felekyan, S.; Gaiduk, A.; Seidel, C. A. M. *J. Phys. Chem. B* **2006**, *110*, 6970–6978.
(68) Kalinin, S.; Felekyan, S.; Valeri, A.; Seidel, C. A. M. *J. Phys. Chem. B* **2008**, *112*, 8361–8374.
(69) Knox, R. S.; van Amerongen, H. *J. Phys. Chem. B* **2002**, *106*, 5289–5293.
(70) Majumdar, Z. K.; Hickerson, R.; Noller, H. F.; Clegg, R. M. *J. Mol. Biol.* **2005**, *351*, 1123–1145.
(71) Haas, E.; Wilchek, M.; Katchalski-Katzir, E.; Steinberg, I. Z. *Proc. Natl. Acad. Sci. U.S.A.* **1975**, *72*, 1807–1811.
(72) Olofsson, M.; Kalinin, S.; Zdunek, J.; Oliveberg, M.; Johansson, L. B. A. *Phys. Chem. Chem. Phys.* **2006**, *8*, 3130–3140.
(73) Kulinski, T.; Visser, A.; Okane, D. J.; Lee, J. *Biochemistry* **1987**, *26*, 540–549.
(74) Borst, J. W.; Laptenok, S. P.; Westphal, A. H.; Kühnemuth, R.; Hornen, H.; Visser, N. V.; Kalinin, S.; Aker, J.; van Hoek, A.; Seidel, C. A. M.; Visser, A. J. W. G. *Biophys. J.* **2008**, *95*, 5399–5411.
(75) Dolghih, E.; Roitberg, A. E.; Krause, J. L. *J. Photochem. Photobiol., A* **2007**, *190*, 321–327.
(76) Dolghih, E.; Ortiz, W.; Kim, S.; Krueger, B. P.; Krause, J. L.; Roitberg, A. E. *J. Phys. Chem. A* **2009**, *113*, 4639–4646.
(77) Di Fiori, N.; Meller, A. *Biophys. J.* **2010**, *98*, 2265–2272.
(78) Kalinin, S.; Valeri, A.; Antonik, M.; Felekyan, S.; Seidel, C. A. M. *J. Phys. Chem. B* **2010**, *114*, 7983–7995.
(79) Maus, M.; Cotlet, M.; Hofkens, J.; Gensch, T.; De Schryver, F. C.; Schaffer, J.; Seidel, C. A. M. *Anal. Chem.* **2001**, *73*, 2078–2086.
(80) Gopich, I. V.; Szabo, A. *J. Phys. Chem. B* **2007**, *111*, 12925–12932.
(81) Isaksson, M.; Norlin, N.; Westlund, P. O.; Johansson, L. B. A. *Phys. Chem. Chem. Phys.* **2007**, *9*, 1941–1951.
(82) Johansson, L. B. A.; Edman, P.; Westlund, P. O. *J. Chem. Phys.* **1996**, *105*, 10896–10904.
(83) Johansson, L. B. A.; Bergstrom, F.; Edman, P.; Grechishnikova, I. V.; Molotkovsky, J. G. *J. Chem. Soc., Faraday Trans.* **1996**, *92*, 1563–1567.